

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 14:51:36 ; Search time 9045 Seconds
(without alignments)
11666.863 Million cell updates/sec

Title: US-09-900-237-29
Perfect score: 3626
Sequence: 1 gcacgaggaccccgctcca.....tcacatttggaggagatttt 3626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_em:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sy:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2595	71.6	3795	8	AF200533	AF200533 Zea mays
2	2595	71.6	3799	6	AX338680	AX338680 Sequence
3	2572.2	70.9	3745	8	AF200528	AF200528 Zea mays
4	2325.2	64.1	3676	8	AF200529	AF200529 Zea mays
5	1699.8	46.9	3723	8	AF150630	AF150630 Gossypium
6	1675.2	46.2	3682	8	AF027174	AF027174 Arabidops
7	1673.6	46.2	3614	6	AX030946	AX030946 Sequence
8	1673.6	46.2	3614	6	AX030960	AX030960 Sequence
9	1408.4	38.8	3725	8	AF200526	AF200526 Zea mays
10	1373	37.9	3752	8	AF200525	AF200525 Zea mays
11	1343.4	37.0	3603	6	AX030942	AX030942 Sequence
12	1343.4	37.0	3603	8	AX030956	AX030956 Sequence
13	1341.8	37.0	3673	6	AX030948	AX030948 Sequence
14	1341.8	37.0	3673	8	AX030962	AX030962 Sequence
15	1340.2	37.0	3812	8	AF200532	AF200532 Zea mays
16	1331.4	36.7	2828	8	AF200527	AF200527 Zea mays
17	1331.4	36.7	2830	6	AX338676	AX338676 Sequence
18	1316.8	36.3	3538	8	AF200530	AF200530 Zea mays
19	1271	35.1	3968	8	AF200531	AF200531 Zea mays
20	1244.8	34.3	3847	8	AF304374	AF304374 Nicotiana
21	1221	33.7	3081	8	AF088917	AF088917 Arabidops
22	1215.6	33.5	3186	8	GHU58283	U58283 Gossypium h
23	1215.6	33.5	3328	6	AR164434	AR164434 Sequence
24	1192.6	32.9	3771	8	AF027173	AF027173 Arabidops
25	1177	32.5	3444	8	AF062485	AF062485 Arabidops
26	1176.6	32.4	3828	6	AX030944	AX030944 Sequence
27	1176.6	32.4	3828	8	AX030958	AX030958 Sequence
28	1171	32.3	87793	2	AP003837	AP003837 Oryza sat
29	1171	32.3	120868	2	AP003748	AP003748 Oryza sat
30	1166.4	32.2	2958	8	AF267742	AF267742 Arabidops
31	1164.2	32.1	3395	8	AF081534	AF081534 Populus x
32	1158.4	31.9	3109	8	AY099636	AY099636 Arabidops
33	1134.2	31.3	3145	8	AF323039	AF323039 Zinnia el
34	1133.6	31.3	3232	8	AF072131	AF072131 Populus t
35	984.2	27.1	2248	6	AX030938	AX030938 Sequence
36	984.2	27.1	2248	8	AX030952	AX030952 Sequence
37	918.2	25.3	2229	8	GHU58284	U58284 Gossypium h
38	746.4	20.6	1529	8	TAU48693	TAU48693 Triticum ae
39	684	18.9	165918	2	AP005420	AP005420 Oryza sat
40	684	18.9	188279	2	AP005579	AP005579 Oryza sat
41	666.2	18.4	7234	6	AX048245	AX048245 Sequence
42	666.2	18.4	7234	8	AF091713	AF091713 Arabidops
43	664.6	18.3	33563	8	AT1086	AT1086 Arabidops
44	648.2	17.9	1515	8	AF323040	AF323040 Zinnia el
45	648	17.9	1699	8	AY059858	AY059858 Arabidops

ALIGNMENTS

RESULT 1
AF200533
LOCUS AF200533
DEFINITION ze mays cellulose synthase-9 (Cesa-9) mRNA, complete cds.
ACCESSION AF200533
VERSION AF200533.1 GI:9622889
KEYWORDS
SOURCE ze mays.
ORGANISM ze mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3795)
Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
Xoconostle-Cazares,B. and Delmer,D.P.

TITLE A comparative analysis of the plant cellulose synthase (CesA) gene family
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3795)
AUTHORS Dhugga, K.S. and Helentjaris, T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES
Location/Qualifiers
1..3795
/organism="Zea mays"
/db_xref="taxon:4577"
/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc."
1..3795
/gene="CesA-9"
238..3477
/gene="CesA-9"
/note="CESA-9"
/codon_start=1
/product="cellulose synthase-9"
/protein_id="AAF89969.1"
/db_xref="GI:9622890"
/translation="MEGDADGVKSGRRGGQVCICGDCGVGTAEAGDFTACDVGFPVCRPYEYERKDGTOACPOCKNKKYKHKGPSALRGEEDDDADDSFNYPASGNDQOKQKATADMRNMRMAGSGDVGRPKYDSGEIGLTKYDSGEIPRGYIPYSYINQSIGEIPGSDTHMSTPTGNIGRRAPFTNMHSSNPFREFSGSVGNVANKERVDGWMKQDKGITPNTGTSIAPSGRGVDDIADSTYNMEDALLNDETQPLSRKVPFPSPRINPYRMYIVLRILVLSIFLHYRITPNPNAYPLWLLSVCEIWFALSDILQDFPKWPIINRET YLDRLALYDREGEPSQLAAVDIFVSTVDMKEPPLVANTVLSILADYPPVDRKVSFYVSDGAAMLTDFALAEISFARKWYFVKKYNIEPRAPWEYFSQKIDYLDKDVHPSFV KDRRAKREYEFKLRVNLVAKAQKVEEGIMODGTWPGNTRDHPGMQVFLGH SGLRTEGNEPLRYVYSREKRPQGHKKKAGAMNALVRVSAVLINGQYMLNDCDHY INNSKALRAMECLMDPNLGRSVICVQFPRFDDGIDRNDRIYANRNTVFEDINLGLDG IQGVPVGTGVFNFTALYGPPIKQKRGFLSSLCGRKKGSKSKSGDGGKSSQKH VDSVPVFNLEDIEEGVEGAGEDDEKSLMSQMSLEKREFGQSAAFVSTIMYGVGPQ SATPESLKEALHVIKSCGYEDKIEWTEIGWIYGSVTEIDILTGFKMHARGWRSIYNT KRPAFKSAPIINSLRNLNOVLWALGSVEILFSRHCPWLVGYGGRULFLERAYINT IYPLSLPLIYCIPLAICLLCKPIELISNFASLWFLSFLSIFATGILBWRSGV GIDEMNQFVWIGISALHFAVQGLKVLAGIDNTNVTVTSKASDEGFAELMYMF KWTLLIPPTIILINLVGVAGISYAINSGYOSWGLPFGKLIFFAFWVLVHLPFLKGLMGRQNRFTIIVVWAILLASIFSLWLRIDPFTNRVTPDTRTCGINC"

BASE COUNT 925 a 927 c 995 g 944 t 4 others
ORIGIN

Query Match 71.6%; Score 2595; DB 8; Length 3795;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;
QY 36 CGGTTTGATCGCTCTCGCGCGCATGCGAGCGACGCGAGCCCTGGAAGTCGGGAGG 95
DB 214 CGGTTTGCGGAGCTCGCTCCATGGAGGCGACGCGAGCGCGTGAAGTCGGGAGG 273
QY 96 CACGGGCGGGACGTGTGCCAGATGTGCCCGACGCGCTGGGACACACCTTTGGACGC 155
DB 274 CGCGGGGAGGCGAGGTGTGCCAGATCTGCGGCGATGCGGTGGGCACTACGCGGAGGGA 333
QY 156 GAGCTCTTACCGCTCGGACGCTGCGGCTTCCCGGTTCGCCGCCCTGCTAGGACAC 215
DB 334 GAGCTCTTACCGCTCGGACGCTGCGGCTTCCCGGTTCGCCGCCCTGCTAGGATAC 393
QY 216 GAGCGAAGGAGGACCCAGGCGCTGCTCCAGTGCAGACCAAGTACAAAGCGCCACAGA 275
DB 394 GAGCGAAGGAGGACACAGCGTGCCTCCAGTGCAGACCAAGTACAAAGCGCCACAG 453
QY 276 GGGAGCCACGATCCCGGGGAGGAGGCGGACACTGATGCCGATGATGTTAGTGAC 335
DB 454 GGGAGTCCAGCGATCCCGGGGAGGAGGAGGACGATGATGCGGATGATGCTAGGAC 513
QY 336 TTCAACTACCTGCATCTGCACTGAGGACGACGAGGAGGATGCTGACAGGATCGC 395

Db 514 TTCAACTACCTGCATCTGCGA 573
QY 396 AGCTGGCGCATGAACACCGGG 455
Db 574 AGCTGGCGCATGAATGCTGGG 633
QY 456 GAGATCGGCTCTCCCAAGTATG 515
Db 634 GAGATCGGCTCTACCAAGTATG 693
QY 516 ACCAACAGCCAGATGTCAGAG 575
Db 694 ACTAACAGCCAGATTTCCGGAG 753
QY 576 CCTACGGGGAACATCAGCAG 635
Db 754 CCTACTGGGAACATTTGGCAGG 813
QY 636 TCAAGGAGTCTCCCGGCAGT 695
Db 814 TCGAGGGAATCTCTGTAGG 873
QY 696 AAAATGAAGCAGCAAGGGTG 755
Db 874 AAAATGAAGCAGCAAGGGGA 933
QY 756 GAAGGTGCGCAGCTACTGAC 815
Db 934 GAGGCGCGGGTGTGTGAT 993
QY 816 CTGAATGATGAACCTCGCCAG 875
Db 994 TTAACCATGAACCTCGCCAG 1053
QY 876 AATCCCTACAGATGGTCAAT 935
Db 1054 AATCCCTACAGATGGTCAAT 1113
QY 936 CGTCTCAAAATCCTGTCGT 995
Db 1114 CGGATCAAAATCCTGTCGT 1173
QY 996 ATTTGGTTGCTTTATCCTGG 1055
Db 1174 ATCTGGTTGCTCTTTCTGG 1233
QY 1056 GAGACCTACCTTGATACAGT 1115
Db 1234 GAGACTTACCTTGATACAGT 1293
QY 1116 GCTGCTGTTGACATATTGTC 1175
Db 1294 GCTGCTGTTGACATTTTGT 1353
QY 1176 GCCAACACTGTGCTATCATT 1235
Db 1354 GCCAATACCTGCTATCATT 1413
QY 1236 GTATCTGATGACGAGGCTTCA 1295
Db 1414 GTATCTGATGAGGCTGCT 1473
QY 1296 GCTAGGAATGGGTACCATT 1355
Db 1474 GCTAGAAATGGGTGCCATT 1533
QY 1356 TACTTTTCCAGAAAATTGAT 1415
Db 1534 TACTTCTCCCAAAAATTGAT 1593
QY 1416 CGCGGGCGCATGAAGAGAGAA 1475

Db 1594 CGCGGGCCATGAAGAGAGAAATATGAAGAATTCAAAATAGGGTAAATGGCCCTTGTGCT 1653
QY 1476 AAGCATTTGAAGTCCCGAGGAGGATGGATCATCAAGATGGCACACCATGGCCAGCA 1535
Db 1654 AAGCCAAAAGTCCCTGAGGAAGATGGATCATCAAGATGGCACACCATGGCCAGCA 1713
QY 1536 AACAATACAGGATCATCCTGGAATGATTCAGGTTTTCCTTGGTGACAGTGGTGCCCTT 1595
Db 1714 AACAATACAGGACCATCCTGGAATGATTCAGGTTTTCCTTGGTGACAGTGGTGCTTT 1773
QY 1596 GATACATGAGGTAATGAGCTCCCGCTTTAGTTTATGTCTGCTGCTGCTGCTGCTGG 1655
Db 1774 GATACATGAGGTAATGAGCTACCCGCTTTGGTCTATGTTTCTGCTGCTGCTGCTGGA 1833
QY 1656 TTCAGCACACAGAGGCTGGTGCCATGAATGCCCTTGTTCGCTGCTCAGCTGCTCTT 1715
Db 1834 TTCAGCATCACAGAAAGCTGGTGCCATGAATGCTTCTGCTGCTGCTGCTGCTGCTT 1893
QY 1716 ACTAATGGCAATACATGTTGAATCTTGATTTGTGATCACTATACATCAACACAGGCT 1775
Db 1894 ACCAATGGCAATACATGTTGAATCTTGATTTGTGATCACTATACATCAACACAGGCT 1953
QY 1776 GTCCGGAAGCTATGCTTCTTAATGGATCCAAAGCTAGGTCGCGAAGTCTGTTATGTG 1835
Db 1954 CTGAGGAAGCTATGCTTCTTATGGATCCCTAAGCTAGGAGGAGTGTCTGCTATGTT 2013
QY 1836 CAGTTCCCAAAAGGTTTGATGGGATTTAGGAATGATCGATATCAACACAGGACACT 1895
Db 2014 CAGTTCCCAAGGTTTCGATGGTATTTAGGAATGATCGATATCCACAGGACACCC 2073
QY 1896 GTCTTTTTCATTAACCTTGGGGCTTGACGGCATTCAGGACACAGTTCATGTTGGA 1955
Db 2074 GTGTTTTCATTAACCTTGGAGGCTTCATGGCATCCAGGACACAGTTCATGTTGGC 2133
QY 1956 ACTGGTGTGTTTCAACAGAACAGTATCTATGTTATGAGCCCCCAATTAAGCGAAG 2015
Db 2134 ACTGGTGTGTTTCAACAGAACAGTCTATATGTTATGAGCCCCCAATTAAGCAAAAG 2193
QY 2016 AAGCCAGGTTCTTGCCATCAGTATGTTGGGCAAGAACAGCAAGTCAAGAA 2075
Db 2194 AAGGGTGTCTTGTCATCAGTATGTTGGGCAAGAACAGCAAGTCAAGAA - 2252
QY 2076 AGGAGCTCAGATAAGAAAAGTGGCAACAGCATGTGGACAGTTCCTTCCAGTATTCAT 2135
Db 2253 --GGGCTCAGACAGAAAAGTCAAGAACATGTGGACAGTTCCTGTCAGTATTCAT 2310
QY 2136 CTCGAGACATAGAGAGGTTTGAAGGTGCTGGGTTGATGATGAGAAATCAGTTCTC 2195
Db 2311 CTTGAAGATATAGAGAGGAGTTGAAGGCGCTGGATTTGATGAGAAATCAGTCTT 2370
QY 2196 ATGCTCAATGAGCTTAGAGAGATTTGGCCAGTACAGCAGCATTTGCTGCCCTCAGT 2255
Db 2371 ATGCTCAATGAGCTTAGAGAGATTTGGCCAACTTCGACGCTTTTGTGCTGCCACT 2430
QY 2256 CTGATGGAATATGTTGTTCTCAGTCTCCACATCCAGAACTCTCTTTGAAAAGAGCT 2315
Db 2431 CTGATGGAATATGTTGTTCTCAGTCTGCGACTCCAGAACTCTCTTGAAGAGCT 2490
QY 2316 ATCCATGCTAATAGTTGTGGCTATGAGGACAGTCTGAATGGGAACTGAGATTTGGTGG 2375
Db 2491 ATCCATGCTAATAGTTGTGGCTACGAGGACAGATTTGAATGGGAACTGAGATTTGGTGG 2550
QY 2376 ATCTATGGAATCTCAGAGAGATATCTTACTGGATTCAGATGACACGACAGGCTGG 2435
Db 2551 ATCTATGTTTCTGTGAGGAAATATCTCAGTGGGTTCAAGATGACACGACAGGCTGG 2610
QY 2436 CGTTCAGTCTATTGCTATGCCAAGCCAGCTTTTCAAGGATCTCCGCCATCAATCTT 2495
Db 2611 CGGTGATCTACTGATGCTTACAGCGCCGCTTCAAGGATCGGCTCCCATCAATCTC 2670
QY 2496 TCAGATGCTGTAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGC 2555
Db 2671 TCAGACGCTCTGAACCAAGTGTCTCGGTGGGCTCTCGGTTCTAGTGGAAATCCTTTTCAGC 2730

QY 2556 CGGCATTGCCCTTATGTTATGGTTACGAGGCGCCCTCAAGTTCTCTGGAGAGATTGCT 2615
Db 2731 CGGCATTGCCCTTATGTTACGGTTACGAGGACGCCCTGAAGTTCTTGGAGAGATTGCC 2790
QY 2616 TACATCAACACACCATTTACCCACTCTCTCCGCTTCTAGTCTATTTGATATATG 2675
Db 2791 TACATCAACACACCATCTACCGCTCAGTCCCTCCGCTCCTCATTTACTGTATCCTG 2850
QY 2676 CCGTCTATCTGCTGCTCAGTGGAAAGTTCATGATCCAGAGATTAGCAACTTGGCCAGT 2735
Db 2851 CCGTGCATCTGCTGCTCAGCGGGAAGTTCATCATCCAGAGATCAGCAACTTCGCTAGT 2910
QY 2736 ATCTGTTTCATTCGCTCTTCTTCAATTTTGGCACTGGTATCTTTCAGATGAGGTGG 2795
Db 2911 ATCTGTTTCATCTCTCTCTTCTCATCTCGATCTTCGCCACGGGTATCTCTGAGATGAGGTGG 2970
QY 2796 AGTGTGTTGGATTCAGAGTGGTGGAGGAATGAACAGTTCCTGGGTCAATTTGGAGGTATC 2855
Db 2971 AGCGGCTGGGCATCGAGAGTGGTGGAGGAACGAGCAGTTCCTGGGTCTATCGGAGCATC 3030
QY 2856 TCTGCACATCTGTTTGGCGCTTTCAGGGTCTTCTGAAGGTGCTTGGCCGCTATCGACACC 2915
Db 3031 TCGCGCCACCTCTTCGCGCTCTTCAGGGCCCTCCTCAAGGTGCTTGGCCGCTATCGACACC 3090
QY 2916 AACTTCACTGTCACTCAAGGCTAATCAGGAAGAGGCGACTTTTGTCTGAGCTCTACATG 2975
Db 3091 AACTTCACTGTCACTCAAGGCTCGGATGAAGAGCGGACTTCGCGGAGCTGTACATG 3150
QY 2976 TTCAAGTGGAGCAGCTTCTCATCCCTCCGACGACATTTTGTATTCATTAACATGTTGGT 3035
Db 3151 TTCAAGTGGAGCAGCTTCTGATCCGCGCCACCACTCTGATCATCACTGCTCGGC 3210
QY 3036 GTGCTTGTGGACCTCTCAGCCATCAACAGTGGTTPACCAATCATGCGGCGCTCTTT 3095
Db 3211 GTTGTGTGGCGCATCTCTACGCCATCAACAGCGGTTACCAGTCTGCGGGTCCGCTCTTC 3270
QY 3096 GGAAGCTCTTCTTGGCTCTTGGGTGATTTTCACTTATACCCATTCCTCAAGGCTCT 3155
Db 3271 GGAAGCTCTTCTTGGCTCTTGGGTGATTTTCACTTACCCGTTACCCGTTCTCAAGGCTCTC 3330
QY 3156 ATGGCAGGCAAAACCGCACACGATTTGTCTGCGGTGCTCTCTCTCTCTCTCTCTCT 3215
Db 3331 ATGGGTCTGGCAGAACCGCACCCGACCATCTGTTGTTGCTGGCGATCTCTCTGCGCTG 3390
QY 3216 ATCTTCTCTCTGCTGTTGGTGGTGGTTCATTCATTCATACCGTCTCTCTCTCTCTCTCT 3275
Db 3391 ATCTTCTCTCTGCTGTTGGTGGTGGTTCATTCATTCATTCATTCATTCATTCATTCAT 3450
QY 3276 ATCCAAACCTGTGGCATCACTGTAGGAAAGTGGGA-GTTTGTAGACACAGAAA 3329
Db 3451 ACTGAACGTGTGGCATCACTGTAGGAGGTGGAAAGTTTGTAGAAACAGAGA 3505

RESULT 2
AX338680
LOCUS AX338680 Sequence 5 from Patent WO0179516.
DEFINITION AX338680
ACCESSION AX338680
VERSION AX338680.1 GI:18129042
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE
AUTHORS Dhugga, K. S. and Helentjaris, T. G.
TITLE Maize cellulose synthases and uses thereof
JOURNAL Patent: WO 0179516-A 5 25-OCT-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1. .3799 source

/organism="Zea mays"
/db_xref="taxon:4577"
238. . 3477
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD20545.1"
/translation="GI:18129043"
/db_xref="GI:18129043"
CDS
636 TCAAGGAGTTCCTCCGCGAGT
637 TCGAGGCAATTCCTGCTAGC
638 TCGAGGCAATTCCTGCTAGC
639 TCGAGGCAATTCCTGCTAGC
640 TCGAGGCAATTCCTGCTAGC
641 TCGAGGCAATTCCTGCTAGC
642 TCGAGGCAATTCCTGCTAGC
643 TCGAGGCAATTCCTGCTAGC
644 TCGAGGCAATTCCTGCTAGC
645 TCGAGGCAATTCCTGCTAGC
646 TCGAGGCAATTCCTGCTAGC
647 TCGAGGCAATTCCTGCTAGC
648 TCGAGGCAATTCCTGCTAGC
649 TCGAGGCAATTCCTGCTAGC
650 TCGAGGCAATTCCTGCTAGC
651 TCGAGGCAATTCCTGCTAGC
652 TCGAGGCAATTCCTGCTAGC
653 TCGAGGCAATTCCTGCTAGC
654 TCGAGGCAATTCCTGCTAGC
655 TCGAGGCAATTCCTGCTAGC
656 TCGAGGCAATTCCTGCTAGC
657 TCGAGGCAATTCCTGCTAGC
658 TCGAGGCAATTCCTGCTAGC
659 TCGAGGCAATTCCTGCTAGC
660 TCGAGGCAATTCCTGCTAGC
661 TCGAGGCAATTCCTGCTAGC
662 TCGAGGCAATTCCTGCTAGC
663 TCGAGGCAATTCCTGCTAGC
664 TCGAGGCAATTCCTGCTAGC
665 TCGAGGCAATTCCTGCTAGC
666 TCGAGGCAATTCCTGCTAGC
667 TCGAGGCAATTCCTGCTAGC
668 TCGAGGCAATTCCTGCTAGC
669 TCGAGGCAATTCCTGCTAGC
670 TCGAGGCAATTCCTGCTAGC
671 TCGAGGCAATTCCTGCTAGC
672 TCGAGGCAATTCCTGCTAGC
673 TCGAGGCAATTCCTGCTAGC
674 TCGAGGCAATTCCTGCTAGC
675 TCGAGGCAATTCCTGCTAGC
676 TCGAGGCAATTCCTGCTAGC
677 TCGAGGCAATTCCTGCTAGC
678 TCGAGGCAATTCCTGCTAGC
679 TCGAGGCAATTCCTGCTAGC
680 TCGAGGCAATTCCTGCTAGC
681 TCGAGGCAATTCCTGCTAGC
682 TCGAGGCAATTCCTGCTAGC
683 TCGAGGCAATTCCTGCTAGC
684 TCGAGGCAATTCCTGCTAGC
685 TCGAGGCAATTCCTGCTAGC
686 TCGAGGCAATTCCTGCTAGC
687 TCGAGGCAATTCCTGCTAGC
688 TCGAGGCAATTCCTGCTAGC
689 TCGAGGCAATTCCTGCTAGC
690 TCGAGGCAATTCCTGCTAGC
691 TCGAGGCAATTCCTGCTAGC
692 TCGAGGCAATTCCTGCTAGC
693 TCGAGGCAATTCCTGCTAGC
694 TCGAGGCAATTCCTGCTAGC
695 TCGAGGCAATTCCTGCTAGC
696 TCGAGGCAATTCCTGCTAGC
697 TCGAGGCAATTCCTGCTAGC
698 TCGAGGCAATTCCTGCTAGC
699 TCGAGGCAATTCCTGCTAGC
700 TCGAGGCAATTCCTGCTAGC
701 TCGAGGCAATTCCTGCTAGC
702 TCGAGGCAATTCCTGCTAGC
703 TCGAGGCAATTCCTGCTAGC
704 TCGAGGCAATTCCTGCTAGC
705 TCGAGGCAATTCCTGCTAGC
706 TCGAGGCAATTCCTGCTAGC
707 TCGAGGCAATTCCTGCTAGC
708 TCGAGGCAATTCCTGCTAGC
709 TCGAGGCAATTCCTGCTAGC
710 TCGAGGCAATTCCTGCTAGC
711 TCGAGGCAATTCCTGCTAGC
712 TCGAGGCAATTCCTGCTAGC
713 TCGAGGCAATTCCTGCTAGC
714 TCGAGGCAATTCCTGCTAGC
715 TCGAGGCAATTCCTGCTAGC
716 TCGAGGCAATTCCTGCTAGC
717 TCGAGGCAATTCCTGCTAGC
718 TCGAGGCAATTCCTGCTAGC
719 TCGAGGCAATTCCTGCTAGC
720 TCGAGGCAATTCCTGCTAGC
721 TCGAGGCAATTCCTGCTAGC
722 TCGAGGCAATTCCTGCTAGC
723 TCGAGGCAATTCCTGCTAGC
724 TCGAGGCAATTCCTGCTAGC
725 TCGAGGCAATTCCTGCTAGC
726 TCGAGGCAATTCCTGCTAGC
727 TCGAGGCAATTCCTGCTAGC
728 TCGAGGCAATTCCTGCTAGC
729 TCGAGGCAATTCCTGCTAGC
730 TCGAGGCAATTCCTGCTAGC
731 TCGAGGCAATTCCTGCTAGC
732 TCGAGGCAATTCCTGCTAGC
733 TCGAGGCAATTCCTGCTAGC
734 TCGAGGCAATTCCTGCTAGC
735 TCGAGGCAATTCCTGCTAGC
736 TCGAGGCAATTCCTGCTAGC
737 TCGAGGCAATTCCTGCTAGC
738 TCGAGGCAATTCCTGCTAGC
739 TCGAGGCAATTCCTGCTAGC
740 TCGAGGCAATTCCTGCTAGC
741 TCGAGGCAATTCCTGCTAGC
742 TCGAGGCAATTCCTGCTAGC
743 TCGAGGCAATTCCTGCTAGC
744 TCGAGGCAATTCCTGCTAGC
745 TCGAGGCAATTCCTGCTAGC
746 TCGAGGCAATTCCTGCTAGC
747 TCGAGGCAATTCCTGCTAGC
748 TCGAGGCAATTCCTGCTAGC
749 TCGAGGCAATTCCTGCTAGC
750 TCGAGGCAATTCCTGCTAGC
751 TCGAGGCAATTCCTGCTAGC
752 TCGAGGCAATTCCTGCTAGC
753 TCGAGGCAATTCCTGCTAGC
754 TCGAGGCAATTCCTGCTAGC
755 TCGAGGCAATTCCTGCTAGC
756 TCGAGGCAATTCCTGCTAGC
757 TCGAGGCAATTCCTGCTAGC
758 TCGAGGCAATTCCTGCTAGC
759 TCGAGGCAATTCCTGCTAGC
760 TCGAGGCAATTCCTGCTAGC
761 TCGAGGCAATTCCTGCTAGC
762 TCGAGGCAATTCCTGCTAGC
763 TCGAGGCAATTCCTGCTAGC
764 TCGAGGCAATTCCTGCTAGC
765 TCGAGGCAATTCCTGCTAGC
766 TCGAGGCAATTCCTGCTAGC
767 TCGAGGCAATTCCTGCTAGC
768 TCGAGGCAATTCCTGCTAGC
769 TCGAGGCAATTCCTGCTAGC
770 TCGAGGCAATTCCTGCTAGC
771 TCGAGGCAATTCCTGCTAGC
772 TCGAGGCAATTCCTGCTAGC
773 TCGAGGCAATTCCTGCTAGC
774 TCGAGGCAATTCCTGCTAGC
775 TCGAGGCAATTCCTGCTAGC
776 TCGAGGCAATTCCTGCTAGC
777 TCGAGGCAATTCCTGCTAGC
778 TCGAGGCAATTCCTGCTAGC
779 TCGAGGCAATTCCTGCTAGC
780 TCGAGGCAATTCCTGCTAGC
781 TCGAGGCAATTCCTGCTAGC
782 TCGAGGCAATTCCTGCTAGC
783 TCGAGGCAATTCCTGCTAGC
784 TCGAGGCAATTCCTGCTAGC
785 TCGAGGCAATTCCTGCTAGC
786 TCGAGGCAATTCCTGCTAGC
787 TCGAGGCAATTCCTGCTAGC
788 TCGAGGCAATTCCTGCTAGC
789 TCGAGGCAATTCCTGCTAGC
790 TCGAGGCAATTCCTGCTAGC
791 TCGAGGCAATTCCTGCTAGC
792 TCGAGGCAATTCCTGCTAGC
793 TCGAGGCAATTCCTGCTAGC
794 TCGAGGCAATTCCTGCTAGC
795 TCGAGGCAATTCCTGCTAGC
796 TCGAGGCAATTCCTGCTAGC
797 TCGAGGCAATTCCTGCTAGC
798 TCGAGGCAATTCCTGCTAGC
799 TCGAGGCAATTCCTGCTAGC
800 TCGAGGCAATTCCTGCTAGC
801 TCGAGGCAATTCCTGCTAGC
802 TCGAGGCAATTCCTGCTAGC
803 TCGAGGCAATTCCTGCTAGC
804 TCGAGGCAATTCCTGCTAGC
805 TCGAGGCAATTCCTGCTAGC
806 TCGAGGCAATTCCTGCTAGC
807 TCGAGGCAATTCCTGCTAGC
808 TCGAGGCAATTCCTGCTAGC
809 TCGAGGCAATTCCTGCTAGC
810 TCGAGGCAATTCCTGCTAGC
811 TCGAGGCAATTCCTGCTAGC
812 TCGAGGCAATTCCTGCTAGC
813 TCGAGGCAATTCCTGCTAGC
814 TCGAGGCAATTCCTGCTAGC
815 TCGAGGCAATTCCTGCTAGC
816 TCGAGGCAATTCCTGCTAGC
817 TCGAGGCAATTCCTGCTAGC
818 TCGAGGCAATTCCTGCTAGC
819 TCGAGGCAATTCCTGCTAGC
820 TCGAGGCAATTCCTGCTAGC
821 TCGAGGCAATTCCTGCTAGC
822 TCGAGGCAATTCCTGCTAGC
823 TCGAGGCAATTCCTGCTAGC
824 TCGAGGCAATTCCTGCTAGC
825 TCGAGGCAATTCCTGCTAGC
826 TCGAGGCAATTCCTGCTAGC
827 TCGAGGCAATTCCTGCTAGC
828 TCGAGGCAATTCCTGCTAGC
829 TCGAGGCAATTCCTGCTAGC
830 TCGAGGCAATTCCTGCTAGC
831 TCGAGGCAATTCCTGCTAGC
832 TCGAGGCAATTCCTGCTAGC
833 TCGAGGCAATTCCTGCTAGC
834 TCGAGGCAATTCCTGCTAGC
835 TCGAGGCAATTCCTGCTAGC
836 TCGAGGCAATTCCTGCTAGC
837 TCGAGGCAATTCCTGCTAGC
838 TCGAGGCAATTCCTGCTAGC
839 TCGAGGCAATTCCTGCTAGC
840 TCGAGGCAATTCCTGCTAGC
841 TCGAGGCAATTCCTGCTAGC
842 TCGAGGCAATTCCTGCTAGC
843 TCGAGGCAATTCCTGCTAGC
844 TCGAGGCAATTCCTGCTAGC
845 TCGAGGCAATTCCTGCTAGC
846 TCGAGGCAATTCCTGCTAGC
847 TCGAGGCAATTCCTGCTAGC
848 TCGAGGCAATTCCTGCTAGC
849 TCGAGGCAATTCCTGCTAGC
850 TCGAGGCAATTCCTGCTAGC
851 TCGAGGCAATTCCTGCTAGC
852 TCGAGGCAATTCCTGCTAGC
853 TCGAGGCAATTCCTGCTAGC
854 TCGAGGCAATTCCTGCTAGC
855 TCGAGGCAATTCCTGCTAGC
856 TCGAGGCAATTCCTGCTAGC
857 TCGAGGCAATTCCTGCTAGC
858 TCGAGGCAATTCCTGCTAGC
859 TCGAGGCAATTCCTGCTAGC
860 TCGAGGCAATTCCTGCTAGC
861 TCGAGGCAATTCCTGCTAGC
862 TCGAGGCAATTCCTGCTAGC
863 TCGAGGCAATTCCTGCTAGC
864 TCGAGGCAATTCCTGCTAGC
865 TCGAGGCAATTCCTGCTAGC
866 TCGAGGCAATTCCTGCTAGC
867 TCGAGGCAATTCCTGCTAGC
868 TCGAGGCAATTCCTGCTAGC
869 TCGAGGCAATTCCTGCTAGC
870 TCGAGGCAATTCCTGCTAGC
871 TCGAGGCAATTCCTGCTAGC
872 TCGAGGCAATTCCTGCTAGC
873 TCGAGGCAATTCCTGCTAGC
874 TCGAGGCAATTCCTGCTAGC
875 TCGAGGCAATTCCTGCTAGC
876 TCGAGGCAATTCCTGCTAGC
877 TCGAGGCAATTCCTGCTAGC
878 TCGAGGCAATTCCTGCTAGC
879 TCGAGGCAATTCCTGCTAGC
880 TCGAGGCAATTCCTGCTAGC
881 TCGAGGCAATTCCTGCTAGC
882 TCGAGGCAATTCCTGCTAGC
883 TCGAGGCAATTCCTGCTAGC
884 TCGAGGCAATTCCTGCTAGC
885 TCGAGGCAATTCCTGCTAGC
886 TCGAGGCAATTCCTGCTAGC
887 TCGAGGCAATTCCTGCTAGC
888 TCGAGGCAATTCCTGCTAGC
889 TCGAGGCAATTCCTGCTAGC
890 TCGAGGCAATTCCTGCTAGC
891 TCGAGGCAATTCCTGCTAGC
892 TCGAGGCAATTCCTGCTAGC
893 TCGAGGCAATTCCTGCTAGC
894 TCGAGGCAATTCCTGCTAGC
895 TCGAGGCAATTCCTGCTAGC
896 TCGAGGCAATTCCTGCTAGC
897 TCGAGGCAATTCCTGCTAGC
898 TCGAGGCAATTCCTGCTAGC
899 TCGAGGCAATTCCTGCTAGC
900 TCGAGGCAATTCCTGCTAGC
901 TCGAGGCAATTCCTGCTAGC
902 TCGAGGCAATTCCTGCTAGC
903 TCGAGGCAATTCCTGCTAGC
904 TCGAGGCAATTCCTGCTAGC
905 TCGAGGCAATTCCTGCTAGC
906 TCGAGGCAATTCCTGCTAGC
907 TCGAGGCAATTCCTGCTAGC
908 TCGAGGCAATTCCTGCTAGC
909 TCGAGGCAATTCCTGCTAGC
910 TCGAGGCAATTCCTGCTAGC
911 TCGAGGCAATTCCTGCTAGC
912 TCGAGGCAATTCCTGCTAGC
913 TCGAGGCAATTCCTGCTAGC
914 TCGAGGCAATTCCTGCTAGC
915 TCGAGGCAATTCCTGCTAGC
916 TCGAGGCAATTCCTGCTAGC
917 TCGAGGCAATTCCTGCTAGC
918 TCGAGGCAATTCCTGCTAGC
919 TCGAGGCAATTCCTGCTAGC
920 TCGAGGCAATTCCTGCTAGC
921 TCGAGGCAATTCCTGCTAGC
922 TCGAGGCAATTCCTGCTAGC
923 TCGAGGCAATTCCTGCTAGC
924 TCGAGGCAATTCCTGCTAGC
925 TCGAGGCAATTCCTGCTAGC
926 TCGAGGCAATTCCTGCTAGC
927 TCGAGGCAATTCCTGCTAGC
928 TCGAGGCAATTCCTGCTAGC
929 TCGAGGCAATTCCTGCTAGC
930 TCGAGGCAATTCCTGCTAGC
931 TCGAGGCAATTCCTGCTAGC
932 TCGAGGCAATTCCTGCTAGC
933 TCGAGGCAATTCCTGCTAGC
934 TCGAGGCAATTCCTGCTAGC
935 TCGAGGCAATTCCTGCTAGC
936 TCGAGGCAATTCCTGCTAGC
937 TCGAGGCAATTCCTGCTAGC
938 TCGAGGCAATTCCTGCTAGC
939 TCGAGGCAATTCCTGCTAGC
940 TCGAGGCAATTCCTGCTAGC
941 TCGAGGCAATTCCTGCTAGC
942 TCGAGGCAATTCCTGCTAGC
943 TCGAGGCAATTCCTGCTAGC
944 TCGAGGCAATTCCTGCTAGC
945 TCGAGGCAATTCCTGCTAGC
946 TCGAGGCAATTCCTGCTAGC
947 TCGAGGCAATTCCTGCTAGC
948 TCGAGGCAATTCCTGCTAGC
949 TCGAGGCAATTCCTGCTAGC
950 TCGAGGCAATTCCTGCTAGC
951 TCGAGGCAATTCCTGCTAGC
952 TCGAGGCAATTCCTGCTAGC
953 TCGAGGCAATTCCTGCTAGC
954 TCGAGGCAATTCCTGCTAGC
955 TCGAGGCAATTCCTGCTAGC
956 TCGAGGCAATTCCTGCTAGC
957 TCGAGGCAATTCCTGCTAGC
958 TCGAGGCAATTCCTGCTAGC
959 TCGAGGCAATTCCTGCTAGC
960 TCGAGGCAATTCCTGCTAGC
961 TCGAGGCAATTCCTGCTAGC
962 TCGAGGCAATTCCTGCTAGC
963 TCGAGGCAATTCCTGCTAGC
964 TCGAGGCAATTCCTGCTAGC
965 TCGAGGCAATTCCTGCTAGC
966 TCGAGGCAATTCCTGCTAGC
967 TCGAGGCAATTCCTGCTAGC
968 TCGAGGCAATTCCTGCTAGC
969 TCGAGGCAATTCCTGCTAGC
970 TCGAGGCAATTCCTGCTAGC
971 TCGAGGCAATTCCTGCTAGC
972 TCGAGGCAATTCCTGCTAGC
973 TCGAGGCAATTCCTGCTAGC
974 TCGAGGCAATTCCTGCTAGC
975 TCGAGGCAATTCCTGCTAGC
976 TCGAGGCAATTCCTGCTAGC
977 TCGAGGCAATTCCTGCTAGC
978 TCGAGGCAATTCCTGCTAGC
979 TCGAGGCAATTCCTGCTAGC
980 TCGAGGCAATTCCTGCTAGC
981 TCGAGGCAATTCCTGCTAGC
982 TCGAGGCAATTCCTGCTAGC
983 TCGAGGCAATTCCTGCTAGC
984 TCGAGGCAATTCCTGCTAGC
985 TCGAGGCAATTCCTGCTAGC
986 TCGAGGCAATTCCTGCTAGC
987 TCGAGGCAATTCCTGCTAGC
988 TCGAGGCAATTCCTGCTAGC
989 TCGAGGCAATTCCTGCTAGC
990 TCGAGGCAATTCCTGCTAGC
991 TCGAGGCAATTCCTGCTAGC
992 TCGAGGCAATTCCTGCTAGC
993 TCGAGGCAATTCCTGCTAGC
994 TCGAGGCAATTCCTGCTAGC
995 TCGAGGCAATTCCTGCTAGC
996 TCGAGGCAATTCCTGCTAGC
997 TCGAGGCAATTCCTGCTAGC
998 TCGAGGCAATTCCTGCTAGC
999 TCGAGGCAATTCCTGCTAGC
1000 TCGAGGCAATTCCTGCTAGC

BASE COUNT 925 a 931 c 995 g 944 t others
ORIGIN
Query Match 71.6%; Score 2595; DB 6; Length 3799;
Best Local Similarity 87.1%; Pred. No. 0;
Matches 2871; Conservative 0; Mismatches 420; Indels 4; Gaps 2;
QY 36 CGGTGGTGGTTCCTGCGCGCATGAGCGGCGAGCGGCGGCGGCGGCGGCGGAGG 95
DB 214 CGGTTCGCGGAGCTCGCGTCCGAGTGGAGGCGGCGGCGGCGGCGGAGTGGGAGG 273
QY 96 CAGCGGCGGCGGAGCTGTCGAGAGTCTGCGCGCGAGCGGCGGCGGCGGCGGCGGAGG 155
DB 274 CCGCGGCGGAGGAGGAGTGTGCCAGATCTCGCGCGGAGTGGCGGCGGCGGCGGAGG 333
QY 156 GAGCTCTTACCGCTTCCGAGCTCTGCGCGTTCGCGGCTGCGCGCGGCGGCGGCGGCGG 215
DB 334 GAGCTCTTACCGCTTCCGAGCTCTGCGCGTTCGCGGCTGCGCGCGGCGGCGGCGGAGT 393
QY 216 GAGCGCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 275
DB 394 GAGCGCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 453
QY 276 GCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 335
DB 454 GCGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 513
QY 336 TCAACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 395
DB 514 TCAACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573
QY 396 AGCTGCGCATGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 455
DB 574 AGCTGCGCATGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 633
QY 456 GAGATCGGCTCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 515
DB 634 GAGATCGGCTCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 693
QY 516 ACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 575
DB 694 ACTAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 753
QY 576 CCTACGGGGAACATCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635
DB 754 CCTACTGGGAACATCAGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 813

GAATGTTCCCTGGAAGAGAGAGAGTTGATGGCTGG 695
GAATGTTCCCTGGAAGAGAGAGAGTTGATGGCTGG 873
TCCCATGACTAATGGGACAAGCATTTGCTCCCTCT 755
TCCCATGACTAATGGGACAAGCATTTGCTCCCTCT 933
TGCATCTACTGAATACAAATGGAAGAGAGCGCTTTA 815
TGCATCACTGATTAACAACATGGAAGAGAGCGCTTTA 993
TATAGAAAAGTCCCATTTGCTTCCCTCCCAATA 875
TATAGAAAAGTCCCATTTGCTTCCCTCCAGGATA 1053
CGGTGTTGTTCTTAAGCATCTTCTCTGCACACTAC 935
ACGATTGATTCTTAAGCATCTTCTCTGCACACTAC 1113
TATCCCACTGTTGGCTTTTATCTGTTATATGTGAG 995
TATCCCACTGTTGGCTTTTATCTGTTATATGTGAG 1173
CGATCAGTCCCGAAGTGGTTTCCCAATCAACCGG 1055
CGATCAGTCCCGAAGTGGTTTCCCAATCAACCGG 1233
AAGGTATGACGAGGAGGTTGAACGCTGTCAGTTG 1115
AAGGTATGACGAGGAGGTTGAACGCTGTCAGTTG 1293
AGTGCACCCCTTGAAGGAGGAGGCTCTCTGTCAC 1175
AGTGCACCCCTTGAAGGAGGAGGCTCTCTGTCAC 1353
AGTGCACCCCTTGAAGGAGGAGGCTCTCTGTCAT 1235
AGTGCACCCCTTGAAGGAGGAGGCTCTCTGTCAT 1413
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1295
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1473
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1355
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1533
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1415
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1593
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1475
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1653
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1535
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1713
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1595
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1773
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1655
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1833
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1715
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1893
GACTTTTGAACGCTTGGCTGAGAGCTTCAAGTTT 1775

/product="cellulose synthase-4"
/protein_id="AAF89964.1"
/db_xref="GI:9622880"

/translation="MEGDADGVKSRGGGVQVCGGVGTAEQDVFAACDVCVGGP
VCPQCEYERKDGTCPCCKTKYKHKGPSAIPRGEEDDADSDNYLNASGNEQK
QKIDMRWRNVNNGSGDVGPRKYDSGEIGLTKYDSGEIPRGIPTVNSQISGEIP
GASPDHMKSPGNTGKRAPFYVNHSPNPSREFSGSIGNVAKRERKQWMDKGT
IPWNGTSIAPSEGRGVGDIDASTYDNMEDALLNDETQPLSRKVPPLSSRINPYRMV
IVLRLLVLFILHYRLTNPVRNAYPLMLLSLVCEIWFALSMILDOFPKWFINRETYL
DRLALYREGPQSQAADVIFSVTPMKPEPLVTANTVLSILAVDIPVKKVSCYVS
DGAAMLTDEALAESEFAKVPFVKYNIIEPRAPWTFQKIDYLDKXVHPSPFXD
RLAMKREYEEFVRVNGLVAKAPPEEGIMODGTPWPNRNRDHPQMIQVFLHSG
DLTEGNEIPRLVYVSRKRPQFQHKKAGAMNALVRVSAVLTQYMLNDCDHYIN
NSKALREACFNLDPNLGRVCYVQPPORFDGIDRNDYANRNTVFEDINLRGLDGIQ
GPVYCTGCVENRTALGYEPPIKQKGFSLSCGGRKKASKSKSDKKSKOKHVD
SSVPENLIEDIEGVGAGFDEKSLMSQNSLEKRFQSSAAFYASTLMELMEYGGVQPSA
TPSLKAEHAIWISCGEDKTEWGTGWIYGVSTEDILITGFKMHARGWVYICMPKR
PAPKGAENLSDRLNQVLWALGVEILFSRCPWYGYGRKLFBRFYAITIY
PLTSIFLLYICLPAICLLTGKFIIPESINFAISWFIISLFIATGILEMRWSGVI
DEWRNEQFWIIGISAHLFVAFQGLKVLAGIDNTFTVSKADEGDFAEIYMKW
TTLIPPTIILINLVGVVAGISYNSGYQSWGFLFKLFFAFWVIYHLVPLKGLM
GRONRPTIVVVAILLASIFSLMVRIDPFTTTRVTPDPTQCGINC"

BASE COUNT 861 a 975 c 1006 g 902 t 1 others
ORIGIN

Query Match 70.9%; Score 2572.2; DB 8; Length 3745;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

Qy	4	CGAGGAACCCGGCTCGTCTGCTCGGTGCGGGTGTGATCGCTCGCGCGCCATGG	53
Db	264	CCAGGCCCCAGGCTCCAGGCGAGCTCCCTCGACGTTCTCGGCGAGCTCGCTGCGCATGG	323
Qy	64	ACGGCGACGGCGCCCTGAAGTCGGGAGGACCGCGGGCGGGGACGTGTGCGCAGATCT	123
Db	324	AGGGCGACGGCGGCTGAAGTCGGGAGGCGGGTGGCGACAGGTGTGCCAGATCT	383
Qy	124	CGCGCGACGGCGTGGGCACACGTTGTGACGGCGACGCTTCTACCGGCTCGCAGCTGTGGC	183
Db	384	CGCGCGACGGCGTGGGCACACCGCGGAGGGGACGCTCTCGCGCGCTCGCAGCTGTGGC	443
Qy	184	GCTTCCGGCTCGCGCCCTGCTACGACGACGACGACGAGGAGGACCCAGGCGCTGCC	243
Db	444	GTTTCCGGTGTGCGCGCCCTGCTACGAGTACGAGGCGAAGGACGACGAGCGGTGCC	503
Qy	244	TCCAGTCAAGACCAAGTACAAGCGCCACAGAGGGAGCCGACGATCCGCGGGAGGAAG	303
Db	504	CCAGTCAAGACCAAGTACAAGCGCCACAGGGGAGCCGCGGATCCCGTGGGAGGAAG	563
Qy	304	CGGACGACGTGATGCGCGATGTAGTACGCTTCAACTACCTGATCTGGCAGTGAAG	363
Db	564	GAGACGACGTGATGCGCGATGTAGTACGCTTCAACTACCTGATCTGGCAATGAGG	617
Qy	364	ACCAGAAGCAAGATGCTGACAGGATGGCGAGCTGGCGCATGAACCCGGGGGCGAGT	423
Db	618	ACCAGAAGCAAGATGCTGCGACAGATGCGCAGCTGGCGCATGAACCTGGGGGCGAGG	677
Qy	424	GCAATGTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAAAGTATGACAGT	483
Db	678	GGGATGTGGTGGCCCAAGTATGACAGTGGCGAGATCGGCGCTTACCAAGTATGACAGT	737
Qy	484	GAGAGATCCCTTAGGGATAGCTCCCTTCACTACCAACACGACGACGATGTCAGGAGAAATCC	543
Db	738	CGGAGATTCCTTCGGGGATACATCCCATCAGTCACTAACAGCCAGATCTCAGGAGAAATCC	797
Qy	544	CTGAGCTTCGCTGTATCATCATGATGTCCCTTACCGGGGAACATCAGCAGACGTGCTC	603
Db	798	CTGCTGTTCCTTCGCTGTATCATCATGATGTCCCTTACCGGGGAACATCAGCAGACGTGCTC	857
Qy	604	CGTTTCCCTATGTGAATCATTCACCAAAATCCGTCAAGGGAGTTCTCCGGCAGTATGGGA	663
Db	858	CATTTCCTATGTGAACCATTCGCCAAATCCCGTCAAGGGAGTTCTCTGTGTAGCATTTGGGA	917

664	ATGTTGCTCGAAAGAGAGAGT	723
918	ATGTTGCTCGAAAGAGAGG	977
724	CCATGACTAATGGGACAGCAT	783
978	CCATGACGAATGGCACAAGCA	1037
784	CATCTACTGAATACAAATGGA	843
1038	CATCAACTGATATCAAAATGGA	1097
844	CTAGAAAAGTCCCATTTGCTTC	903
1098	CTAGAAAAGTCCCATTTGCTTC	1157
904	GTTTGGTGTCTTAAGCATCTT	963
1158	GATTGATGTCTTAAGCATCTT	1217
964	ACCCACTGTGGCTTTTATCTGT	1023
1218	ACCCATATGCTTCTATCTGT	1277
1024	ATCAGTTCCCGAAGTGGTTTCC	1083
1278	ATCAGTTCCCTTAAGTGGTTTCC	1337
1084	GGTATGACCGAAGAGTGAACCC	1143
1338	GGTATGACCGGGAAGGTGAGCC	1397
1144	TCGACCCCTTTGAAGGAGCCACC	1203
1398	TCGACCCCAATGAAGAGCCACC	1457
1204	TTGATTTATCCGTGGGACAGGT	1263
1458	TGGATTTACCTTGGTGAAGGT	1517
1264	CTTTTGGACGATGGCTGAGAC	1323
1518	CATTGTGACACTAGCTGAGAC	1577
1324	AGTATGACATTTGAACCCAGAC	1383
1578	AGTACAACATTTGAACCTAGAC	1637
1384	AAGACAAGTCCAGCCTTCATT	1443
1638	AGGACAAGTGCACCCCTTCATT	1697
1444	AATTTAAATCAGGATAAATGC	1503
1698	AATTCAGATTTAGGTAATGGA	1757
1504	GGATATGCAAGATGGCACACC	1563
1758	GGATATGCAAGATGGCACACC	1817
1564	TTCAAGTGTTCCTTGGTTCACAG	1623
1818	TTCAAGTGTTCCTTGGTTCACAG	1877
1624	TAGTTTATGTCTCGTGAAAG	1683
1878	TGGTCTATGTCTCGTGAAAG	1937
1684	TGAATGCCCTTGTGCTGCTCA	1743
1938	TGAATGCCCTTGTGCTGCTCA	1997
1744	ATTGTGATCTACATCAACAAG	1803

Db 1998 ATTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGATATGTGCTTCCTTATGG 2057
Qy 1804 ATCCAAACCTAGTCCCAAGTCTGTATTATGTGCAGTTCACCAAGGTTTGTATGGGATG 1863
Db 2058 ACCCTAACCTAGGAAGAGTGTCTGTACGTCCAGTTTCCCAAGAGATTCCGATGGCATTG 2117
Qy 1864 ATAGGAATGATCGATATGCAAAACAGNACACTGTCTTTTGTATATTAATCTTGAGGGCC 1923
Db 2118 ACAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTTTCGATATTAATCTTGAGAGTC 2177
Qy 1924 TTGACGCACTTCAAGGACCAAGTTATGTGGAACTGGTGTCTTTTCAACAGAACAGCTA 1983
Db 2178 TTGATGGCATCCAAAGGACCAAGTTATGTGCGAACTGGCTGTCTTTCAACCGCAAGCTC 2237
Qy 1984 TCTATGTTTATGAGCCCCCAATTAAGCGGAAGACCAAGTTCCTTGGCATCACTATGG 2043
Db 2238 TATATGTTTATGAGCCCCCAATTAAGCAGAAAGGTTGTTCTTGTCACTACTATGG 2297
Qy 2044 GGGCAAGAAAGGCAAGCAAGTCAAGAAAGGAGGCTCAGATAGAAAAGTCAACA 2103
Db 2298 GCGGTAGGAAGAGGCAAGCAAAATCAAAAGAA---GGGCTCGGCAAGAAAGTCCGAGA 2354
Qy 2104 AGCATGTGGACAGTCTGTTCACAGTATTCAATCTCGAAGACATAGAGGAGGTTTGAAG 2163
Db 2355 AGCATGTGGACAGTCTGTTCACAGTATTCAACCTTGAAGATATAGAGGAGGAGTTGAAG 2414
Qy 2164 GTGCTGGGTTTGTATGATGAGAAATCAGTCTCATGTCTCAATAGCTTAGAAGAGAT 2223
Db 2415 GCGTGGATTTGACGAGAGAAATCACTCTTATGCTCAAAATGAGCCTGGAGAAGAT 2474
Qy 2224 TTGGCCAGTCAGACAGATTTGTTGCCCTCCACTCTGATGGAATATGTTGGTTCTTCAGT 2283
Db 2475 TTGGCCAGTCGACAGCTTTGTTGCCCTCCACTCTGATGAGATATGTTGGTTCTTCAGT 2534
Qy 2284 CCTCCATCCAGAAATCTTTTGAAGAAGCTATGCCATGTATCAAGTTGTGGCTATGAGG 2343
Db 2535 CCGCAACTCCGGAGTCTCTTCTGAAGAAGCTATCCATGTATATAAGTGTGGCTATGAGG 2594
Qy 2344 ACAAGTCTGAATGGGAACTGAGATTTGGTGGATCTATGATCTGTACAGAGATATTC 2403
Db 2595 ACAAGTCTGAATGGGAACTGAGATTTGGTGGATCTACGGTCTGTGACAGAGATTC 2654
Qy 2404 TTACTGGATCAAGATCAGCAGAGGCTGGGTTTCAAGTCTATTGTCATGCCCAAGCGCC 2463
Db 2655 TCACCGGATCAAGATCAGCGGAGGCTGGGTTGCTACTGATCTGATGCCCAAGCGCC 2714
Qy 2464 CAGCTTTCAAGGATCTGCCCATCAATCTTTTCAGATCTGTGAACCAAGTCTCGGT 2523
Db 2715 CAGCTTTCAAGGATCTGCCCATCAATCTTTTCGAGCCTGTGAACCAAGTCTCGGT 2774
Qy 2524 GGGCTCTCGGTTCTGTTGAATCTTTTACGGCGCAATGCCCTTATGATGGCTACG 2583
Db 2775 GGGCTCTGGTCTCGTGAGATCTCTTACGGCGCACTGCCCTTGTGATGGCTACG 2834
Qy 2584 GAGGGCGCTCAAGTCTCGTGAGAGATTCGCTTACATCAACACCACTTACCCACTAA 2643
Db 2835 GAGGGCGCTCAAGTCTCGTGAGAGATTCGCTTACATCAACACCACTTACCCACTCA 2894
Qy 2644 CCTCTCCCGCTCTAGTCTATGTATATGCTGTATCTGTCTGTCACACCTGGAAGT 2703
Db 2895 GGTCCATCCCGCTCTCTACTGTGATCTGCTGCGCGCATCTGTCTGCTCACCGGAAGT 2954
Qy 2704 TCATCATGCCAGATTTAGCAACTTGGCAGTATGCTGTTTCATTTGGCTCTTCTTCA 2763
Db 2955 TCATCATGCCAGATTCAGCAACTTCCAGCATCTGTTTCATCTCCCTCTTCTCATCTCGA 3014
Qy 2764 TTTTCGCCACTGTTATCTTTCAGATAGGTGGAGTGGTGTGGCATTTGACGAGTGGTGA 2823
Db 3015 TCTTCGCCACGGCATCTCGAGATAGGTGGAGCGGGGTGGCATCGACGAGTGGTGA 3074
Qy 2824 GGAATGAACAGTCTTGGGTCTATTTGGAGGTATCTTCACACATCTTGTTCGGTCTTTCAG 2883

Db 3075 GGAACGACGAGTTCGTGGTGATCGGGGCGCATCTCCGGCGACCTCTTCGCCCGTGTTCAGG 3134
Qy 2884 GTCTTCTGAAGTGTGTTCCGGTATCGACACCAACTTCACGTGTACCTTCAAGGCTAATG 2943
Db 3135 GCCTGCTCAAGTGTGTCGGCGCATCGACACCAACTTCACCGTCACTTCCCAAGGCTCGG 3194
Qy 2944 AGGAAGAAGGCGACTTTCCTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCCTC 3003
Db 3195 ACGAGGACGGGAGCTTCGCGGAGCTGTACATGTTCAAGTGGACGACGCTTCTGATCCCGC 3254
Qy 3004 CGACACCATTTTGTATCATTAACATGTTGCTGCTGCTGGCGACCTCTTACGCCATCA 3063
Db 3255 CCACACCATCTGTATCATCAACCTGGTTCGGGCTGCTGCCGCGCATCTCTTACGCCATCA 3314
Qy 3064 ACAGTGGTTACCAATCATGCGGGGCGCTCTTTTGGGAAGCTCTTCTTGGCTTCTGGGTGA 3123
Db 3315 ACAGCGGATACGATCGTGGGCGCGCTCTTTCGGCAAGCTCTTCTGCGCTTCTGGGTCA 3374
Qy 3124 TTGTTCACTTATACCATTCCTCAAGGCTTATGCGGCGCAAAACCGCACACGACGA 3183
Db 3375 TCGTCCACTGTACCGTTCCTCAAGGCGCTCATGGGCGAGCAGAACCGCACCGACCA 3434
Qy 3184 TTGTCATGCTGGGCTGTCTCTCGCTTCTATCTTCTGCTGTGGGTTCTGTGTTG 3243
Db 3435 TCGTGTGCTGTGGGCGCATCTGCTGGCGTCCATCTTCTTCTGCTGTGGGTTCCGATCG 3494
Qy 3244 ATCCATTCACATCCCGTCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACTGCTAGG 3303
Db 3495 ACCCTTTCACACCGCGCTCACTGGCCGGATACCGACGCTGGCATCAACTGCTAGG 3554
Qy 3304 AAAGTGGAGTTT 3316
Db 3555 GAAGTGGAGGTT 3567

RESULT 4
AF200529
LOCUS Zea mays cellulose synthase-5 (Cesa-5) mRNA linear PLN 31-AUG-2000
DEFINITION Zea mays cellulose synthase-5 (Cesa-5) mRNA, complete cds.
ACCESSION AF200529
VERSION AF200529.1 GI:9622881
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1 (bases 1 to 3676)
AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,
Xoconostle-Cazares,B. and Delmer,D.P.
TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene family
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3676)
AUTHORS Dhugga,K.S. and Helentjaris,T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES
source Location/Qualifiers
1..3676
/organism="Zea mays"
/db_xref="taxon:4577"
/note="identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc"
1..3676
/gene="Cesa-5"
254..3484
/gene="Cesa-5"
/note="Cesa-5"
/codon_start=1

Query Match	64.1%;	Score 2325.2;	DB 8;	Length 3676;
Best Local Similarity	82.8%;	Pred. No. 0;		
Matches 2706;	Conservative	0;	Mismatches 549;	Indels 12;
				Gaps 4;

Qy	698	AATGAACAGGACAAAGGGTGC	GCATGACTAATCGGACAAAGCATTTGCTCCCTCCTGCA	751
Db	886	AATGA---AGGATAAAGTGCA	CTATGACCAATGGAACAAAGCATTTGCTCCATCAGA	942
Qy	758	AGGTGGCGAGCTACTTGACATC	CATCTACTGAATACAACTATGGAAGACGCTTTACT	817
Db	943	AGGGCGTGAGTTGCTGATATT	CTTCTACTGATTATACATGGAAGATGCTTTACT	1002
Qy	818	GAATGATGAACATCGCCAGCCT	CTAGAAAAGTCCCAATTTGCTTCTCCCAAAATAAA	877
Db	1003	GAATGATGAACATCGGCAACCT	CTAGAAAAGTGCCAATTTCTTTCATCCAGAATAAA	1062
Qy	878	TCCTTACAGNATGGTCATTGTT	SGTTGGTCTTCTTAAGCATCTTCTCGTCACTACCG	937
Db	1063	TCGGTACAGNATGGTCATTGTT	STTTGGCTGTTCTATGTCATATTTCTTTCGCGCTACCG	1122
Qy	938	TCTCACAAATTCCTGTGCGTAAT	ACCCACTGTGGCTTTTATCTGTTATATGTGAGAT	997
Db	1123	TATCACACATCCGTGCAACAA	ATCCACTGTGGCTTTTATCCGTCATATGTGAGAT	1182
Qy	998	TTGGTTTGCTTTATTCCTGGATG	ATCAGTTCCGGAAGTGGTTTCCAATCAACCGGGA	1057
Db	1183	CTGGTTTGCTTTGTCTCTGGATT	ATCAGTTCCCAAGTGGTCCCAATCAACCGGTGA	1242
Qy	1058	GACCTACCTTGATAGACTGGCT	SGTATGACGCGAAGGTGAACCGTCTCAGTTGGC	1117
Db	1243	AACATACCTTGATAGACTGGCT	SGTATGACGCGAAGGTGNAACCATCTCAATTAGC	1302
Qy	1118	TGCTGTGTACATATTTGTCCAGT	TCGACCCCTTGAAGGAGGCACTATCGTCACTGC	1177
Db	1303	TCCTGTGTGATATTTTGTCCAGT	TGGATCCAATGAAGGAGCGCTCTCTTTGTCACTGC	1362
Qy	1178	CAACACTGTGCTATCCATTCTTT	TTGATTATCCCTGGACAGAGTCTCTTGGCTATGT	1237
Db	1363	AAATACTGTGCTTTTCCATCCTTT	TCGATTATCCGGTTGACAGGTATCTTGTCTATGT	1422
Qy	1238	ATCTGATGACGGAGCTTTCAATG	CTTTTGACGCATTTGGCTGAGACTTCAGAGTTTGC	1297
Db	1423	TTCGGATGATGAGCTGCTATG	CTTTTGATGCTCTCTCTGAAACTTCAGAGTTTGC	1482
Qy	1298	TAGAAATGGGTACCATTGTTG	AGTATGACATTGAAACCCAGAGCTCCCGAGTTTGA	1357
Db	1483	TAGAAATGGGTTCCGTTCTGT	AGTACAACTATAGAGCTANGGCCCGGAATGGTA	1542
Qy	1358	CTTTTGCCAGAAAATTGATTAG	AAGACAAAAGTCCAGAGCTTCATTTTGTTAAAGACCG	1417
Db	1543	CTTTTGCTCAGAAAATTGATTAG	AAGACAAAAGTTCAAACCTCATTTGTGAAAGAACG	1602
Qy	1418	CCGGGCCATGAAGAGAGATATG	AATTTAAATCAGGATAAATGCCCTAGTTTCTAA	1477
Db	1603	CCGGGCCATGAAGAGAGATAAT	AATTTCAAAGTTCGTATCAATGGCTTTGTAGCCAA	1652
Qy	1478	GGCATTGAAAGTCCCGAGGAAT	GGATCATGCAAGATGGCACACCATGCCCAGGAAA	1537
Db	1663	GGCACAAAAGTTCCCGAGGAG	GGATCATGCAAGATGGTACACCTTGGCCTGGGAA	1722
Qy	1538	CAATACAGGGATCATCCTGGGA	TTTCAGGTTTTCTTGGTCAAGTGGTGGGCTTTGA	1597
Db	1723	CAATACTAGGACCATCCTGGG	TTTCAGGTTTTCTTGGGTCACAGTGGGAGGGCTTGA	1782
Qy	1598	TACTGAGGTTAATGAGCTCCCG	TAGTTTTATCTGTCTCGTGAAGAACGCTCTGGGTT	1657
Db	1783	CGTTGAAGGCAATGAACCTCCG	TGGTTATGTGTCTCGTGAAGAACGCTCTGGNAT	1842
Qy	1658	CCAGCACCAAGAAGGCTGGT	TGAATGCCCTTGTTCGTGTCTCAGCTGTCTCTTAC	1717
Db	1843	CCAACATCACAGAAGCGTGGT	TGAATGCACCTGTTTCGTGTATCAGCTGTCTCTTAC	1902
Qy	1718	TAATGGCAATAATGTTGAAAT	ATTGTGATCACTATACATACACACAGCAGGCTGT	1777
Db	1903	TAATGGCAATAATGTTGAAAT	ATTGTGACCACTACATCAATATAATACAGAGGCTCT	1962
Qy	1778	CCGAGAGACTATGTGCTTCTCT	ATCCAAAACCTAGTTCGGCAAGTCTGTATTGTGCA	1837

Db	1963	TCGAGAAAGCTATGTGCTTCCTTATTCGGACCCAAACCTAGGAAGGAATGCTGTGTATGTCCA	2022
Qy	1838	GTTCACCAAAAGTTTGATGGGATTTAGAGAAATGATCGATATGCAACACAGGACACTGT	1897
Db	2023	ATTTCCTCAGAGGTTTGATGGTATTGATAGGAATGACCGATATGCAACACAGGAACACTGT	2082
Qy	1898	CTTTTTTGATATTAACCTTGAGGGGCTTGAGGGCATTCGAAGACACAGTTATGTGGGAAC	1957
Db	2083	GTTTTCGATATTAACCTTGAGAGGCTTGACGGCATTCGAAGGCCAGTTTATGTGGGAAC	2142
Qy	1958	TGTTTGCTGTTTTCACAGAACAGCTATCATGTAGTTTATGAGCCCCCAATTAAGGGGAAGAA	2017
Db	2143	TGTTTGCTGTTTAAACAGAACGGCTTATATGGTTATGAGGCTCCAGCTCAGAGAAAAAAA	2202
Qy	2018	GCCAGGTTTCTTGGCATCACTATGTGGGGCAAGAAGAAGCAAGCAAGTCAAGAAGAAAG	2077
Db	2203	GCCAGGCTTCTTCTCTTCGCTTTGGGGGAAGAAAAGACGTCAAAATCTAAGAA---	2259
Qy	2078	GAGCTCAGATAAGAAAAAGTCGAACAAGCATGTGGACAGTTCTGTTCACAGTATTCAATCT	2137
Db	2260	GAGCTCGGAAAAGAAGACTCACATAGACACGACACAGTTCTGTACCAGTATTTAATCT	2319
Qy	2138	CGAAGACATAGAGGAGGTTTGAAGGTGCTGGGTTTCATGATGAGAAAAATCAGTTCAT	2197
Db	2320	CGAAGATATAGAGGAAGGATTTGAAGGTTCTCAGTTTGATGATGAGAAATCGCTGATTAT	2379
Qy	2198	GTCTCAAAATGAGCTTAGAGAAGAGATTTGGCCAGTCAGCAGCATTTGTTGCCTCCACTCT	2257
Db	2380	GTCTCAAAATGAGCTTGAGAGAAGATTTGGCCAGTCAGTGTGTTTTGAGGCTCTACTCT	2439
Qy	2258	GATGGAATATGGTGGTTCCTCAGTCCTCCACTCCAGAAATCTCTTTTGAAGAAGCTAT	2317
Db	2440	GATGGAATATGGTGGTTCCTCCAAATCTCGAACTCCAGAGTCTCTCTTGAAGAAGCTAT	2499
Qy	2318	CCATGTCATAAGTTTGGCTATGAGGACAAGTCTCAATGGGGAACATGAGATTGTTGGAT	2377
Db	2500	TCATGTGTCATGCTGGCTATGAGGACAAAACTGACTGGGGAACATGAGATTGGGTGGAT	2559
Qy	2378	CTATGGATCTGTCCACAAGAGATTTCTTACTGGATTCAAGATGCAAGCAAGAGGCTGGCG	2437
Db	2560	CTATGGTTCGTTACAGAAGACATCTCACCGGATTTCAAGATGCATGCTCGAGGCTGGCG	2619
Qy	2438	TTCAGTCTATTTGCATCCCAAGCGCCCACTTTCGAAGGGATCTGCCCCATCAATCTTTC	2497
Db	2620	ATCAATCTACTGCATGCCTTAAGCGACAGCTTTCAAGGATCTGCTCTCATCAACCTTTC	2679
Qy	2498	AGATCGTCTGAACCAAGTGTGGGTGGGCTCTCGGTTCTGTTGAAATTTCTTTTCAGCGG	2557
Db	2680	GGATCGTTTGAATCAAGTGTCTGGTGGGCTCTTGTTCCATTTGAAATTTCTTTTCAGCAG	2739
Qy	2558	GCATTTGCCCTTATGTATGGTACGAGGGGCGCTCAAGTTTCTGTGAGAGATTCGCTTA	2617
Db	2740	GCATTTGCCCATATGGTATGGCTATGGAGGGCGGCTTAAATTCCTGGAGAGATTGCTTA	2799
Qy	2618	CATCAACACCACCATTTACCCATACCTCTCTCCCGTCTTAGTCTATTGTATATTGCC	2677
Db	2800	TATCAACACACAATTTATCCACTCACATCAATCCGCTCCTCCTGTACTGCATATTGCC	2859
Qy	2678	TGCTATCTGCTGCTCAGCTGGAAAGTTCAATCATGCGCAGAGATTAGCAACTTTGGCCAGTAT	2737
Db	2860	AGCAGTTTGTCTTCTCACTGGGAAGTTCAATCATCCAAAGATTAGTAACCTTAGAGAGTGT	2919
Qy	2738	CTGGTTCAATTCGCTCTTCTTCAATTTTCCCACTGGTATCCCTGAGATGAGGTGGAG	2797
Db	2920	TTGGTTTATATFCGCTCTTATCTCAATTTTGCCACTTGGTATTCCTTGAGATGAGGTGGAG	2979
Qy	2798	TGTTGTTGGCATTTCAGCAGTGGTGGAGGAATGAACAGTTTCTGGGTCAATTTGGAGGTATCTC	2857
Db	2980	TGGTTGGCAATTTGATGAATGGTGGAGGAACAGCAGTTCTGGGTCAATTTGGTGTATTTC	3039
Qy	2858	TGCACATCTGTTTGCCTCTTTCTAGGGTCTTCTGAAGGTGCTTGCCTGATTCGACACCAA	2917

Db	3040	TGCGCATTTATTTGCCCGTCTTCCAGGGTCTCCTGAAAGGTGCTTGTGATATCGACACGAG	3099
QY	2918	CTTCACGTGTACCTCAAGGCTAATGACGAAGAAGGGACATTTGTGCTGAGCTCTACATGTT	2977
Db	3100	CTTCACGTGTACCTCTAAGGCCACTGACGAAGAAGGTGATTTTCCGAGACTCTACATGTT	3159
QY	2978	CAAGTGGACGACGTTCTTCATPCCCTCGGACGACCATTTTGATCATTAACATGGTTGGTGT	3037
Db	3160	CAAGTGGACAACGCTCTGTGATCCCAACCAACCATTTTGTGATCATCAACCTGGTCGGCGT	3219
QY	3038	CGTTGCTGGCACCTCCCTACGCCATCAACAGTGGTTACCAATCATGGGGGCCCTCTTTGG	3097
Db	3220	GGTGCCTGGCATTTCCCTACGCAATCAATAGCGGTTACACGATGAGGGACCTCTTTTCGG	3279
QY	3098	GAAGCTCTCTTTGCGCTTCTGGGTGATTTGCTCACTTATACCATCTCTCAAGGGTCTTAT	3157
Db	3280	GAAGCTCTCTTTGGGTCTGGGTGATTTGCTCACTTATACCATCTCTCAAGGGTCTTAT	3339
QY	3158	GGGAGGCAAAACCGACACCGACGATTTGTCATCGTCTGGGCTGTCTCTTCGCTTCTAT	3217
Db	3340	GGGAAGCAGAACCGACGACGATTTGTCATCGTCTGGGCTGTCTCTTCGCTTCTAT	3399
QY	3218	CTTCTCTCTGCTGGGTTCTGTTGATTCATTCACCTACCGCTCTGCTGCGCCCAATAT	3277
Db	3400	CTTTTCCCTGATGGGTTCTGATCGATTCACCTACCGCTCTGCTGCGCCCAATAT	3459
QY	3278	CCAAACCTGTGGCATCAACTGCTAGGA	3304
Db	3460	CGCGAAATGTGGCATCAACTGCTAGGA	3486
RESULT 5			
AF150630			
LOCUS			
DEFINITION	AF150630	Gossypium hirsutum cellulose synthase catalytic subunit (celA3)	
ACCESSION	AF150630	mRNA, complete cds.	
VERSION	AF150630.2	GI:6446576	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			
FEATURES			
source			

gene /dev_stage="24 days post anthesis"
1..3723
/gene="celA3"
CDS 269..3472
/gene="celA3"
/note="celA1 homolog"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="A039534.2"
/db_xref="GI:6446577"
/translation="MESEGDIGKPMKNGGQTCQICGDNVGNKNTDGPFIACNICAF
PVRCPCYEYERKDNQSPQCKRYKWKSPAILRGDRETDGADGASDFISENQE
QOKRLAERGWNKAYGRGDEGAPTYDKESHNIHPIPLTSGOEVSGLSAAASPERLS
MASPGVAKSSIRVDPVREFSGSLGNVAKWERYDKWKKOEKNTVPMSTCOATSE
RGLGDIDASTDVLDSOLNDEARPLSRKVSYSKINPYRMVILRLVILFLHY
RINYPNNAWLHLISVICIWFALISWILDQPKWLPVNRETYLRLALRYDREGEPS
ELAAVDIFSVTDPLKEPLVNTANVLSILAVDPVKVSCYVSDDGAAWLTEPAUSE
TSFARKQVPEGMIMQDTPWPNNTDHPGMIQVFLGSGDLDAEGNELPRVY
VSRKRFQGHKKAGAMNALVRVSAVLTNGAFLLNDCDHYINNSKALREAMCFLLMD
INGLVAKOVRPEGMIMQDTPWPNNTDHPGMIQVFLGSGDLDAEGNELPRVY
PNLGKQVYVQFQRFQDIDRNDYANRNTVFFDINLRGLDGIQGVYVGTGVFNRT
ALYGEPLAPKPKRTGILSLGSGSRKSSKSKGSDKKSGKHVDSPVPVFNED
IERGVEGAGDEDEKSLMSOMSELEKRFQSAVAVSTLMENGVPQSAPEPILLLKEAL
HVSQGEDTDMGSEIGWYGSVTEIDILGFRKHARGWRSIYCMKRPAPKAPSLPIN
LSRLNOVLWALGSVEILFSRHCPIWYSGRLKWLREFAVYNTTPIYPTALPLMY
CTLPVCLLNFKFIIPQISNLASIFLSIFATGILKMKWNGVIGDQWNRNEQFW
VIGVSAHLFAVFQGLKVLGADIDNFTVTSKASDEGDFAEIYMKPWTLLIPPTTL
LIINLLVWAGISVIVNSGYQSMPLGKLFKFAFWIILHYPLFKLGLMGRQNRPTIV
VWNSILLASIFSLWLVRIDPFTTRVTPDVEQGINC"
BASE COUNT 1029 a 700 c 880 g 1114 t
ORIGIN

Query Match 46.9%; Score 1699.8; DB 8; Length 3723;
Best Local Similarity 72.6%; Pred. No. 0;
Matches 2346; Conservative 0; Mismatches 827; Indels 60; Gaps 9;
QY 84 AAGTCCGGAGGACGCGGGACGCTGCCAGATCTCGCGGACGGGCTGGGCACC 143
DB 296 AAGCCCATGAGAATCTTGGTGCCGAGACATGCCAATCTGTGTGACATGTTGGCAAA 355
QY 144 ACGTTGGACGGCAGCTTCCACGCCCTGGACGCTGTCGCCCTTCCCGGCTGCGCGCCC 203
DB 356 AATACAGATGGTGATCCATTCTGTCGAATATATGTGATCTTCTGTATCGAGGCT 415
QY 204 TGCTACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 263
DB 416 TGTTATGAATATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 475
QY 264 AAGCGCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 323
DB 476 AAGTGGCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 535
QY 324 GATGCTAGTACTCACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 383
DB 536 GATG---GCGCCAGTGATTTCTTACCTCAGAAACCAAGAGCAGACGACGACGACGAC 592
QY 384 GACAGGATGCGCAGCTGCGCATGAACACCGGGGCGAGTGCAATGTTGGCCACCCCAAG 443
DB 593 GAAAGATGCAAGGCTGGAATGCAAGATGCTAGAGGGAGGAGTGTGCTGCCACT 652
QY 444 TATGACAGTGGGATCGGCTCTCCAAATGATGACAGTGGAGAGATCCCTTAGGGGATAC 503
DB 653 TAT-----GATAAAGAAATTCATAAACCATATTCCTGCTCACCAGTGGGCGAG 703
QY 504 GTCCCTTCAGTCACCAACAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCCCTGATCAT 563
DB 704 G-----AGGTTCTGTTGAGCTATCTGCTGCACTACCTGAGCGG 742
QY 564 CACATGATGTCCTTACCGGGAACATCAGCAGACGCTGCTCGCTTCCCTATGTGAATCAT 623
DB 743 CTTTCTATGGCATCTCCTGG-----AGTTGCTGGTGGGAATCGAGTATAGGGTGT 795
QY 624 TCACCAAAATCCGTCGAAGGAGGATCTCCGGCAGATATGGGAATGTTGCTCGGAAAGAGAGA 683

DB 796 GGATCCAGTCAGGAGGATTTGG 853
QY 684 GTTGATGGCTGGAAAAATGAAG 743
DB 854 GTTGATGGCTGGGAAGATGAAG 907
QY 744 ATTGCTCCCTCTGAAGTCGG 803
DB 908 CAGGCTACTTCTGAA---AGAG 964
QY 804 GAAGACGCTTTTACTGTAATGAT 863
DB 965 GATGACTCTCAACTGAATGAT 1024
QY 864 TCCTCCAAATAAATCCCTAC 923
DB 1025 TCATCTAAAATAAATCCCTAT 1084
QY 924 TTCTGCACTACCGTCTTCACA 983
DB 1085 TTCTTGCACTACCGAATAACA 1144
QY 984 GTTATATGTGAGATTTGGTTG 1043
DB 1145 GTGATATGTGAGATTTGGTTG 1204
QY 1044 CCAATCAACCGGAGACCTAC 1103
DB 1205 CTTGTTAATCGTGAGGATAC 1264
QY 1104 CCGTCTCAGTTGGCTGCTGTTG 1163
DB 1265 CCGTCCGAATTAGCTGCAGTTG 1324
QY 1164 CCTATCGTCACTGCCAACACTG 1223
DB 1325 CCACCTTGTGACGCCACACTG 1384
QY 1224 GTCTCTTGTGCTATGTATCTGAT 1283
DB 1385 GTCTCTTGTGCTATGTTCTGATG 1444
QY 1284 ACTTCAGAGTTTGTAGTAAAT 1343
DB 1445 ACATCAGAGTTTGCAGGAAAT 1504
QY 1344 GCTCCCGAGTTTACTTTTGC 1403
DB 1505 GCTCCAGAAATGGTATTTTGCAC 1564
QY 1404 TTTGTTAAAGACCCCGCGGCA 1463
DB 1565 TTTGTTAAAGATCGTAGAGT 1624
QY 1464 GCCCTAGTTTCTAAGGCATTGA 1523
DB 1625 GGGCTTCTCGCAAGGACACAG 1684
QY 1524 CCATGGCCAGGAACAATACCA 1583
DB 1685 CCATGGCTGTGATAACACCA 1744
QY 1584 AGTGGTGGCTTGTACTACTGAG 1643
DB 1745 AGTGGGGCCCTGTGATGCTGAG 1804
QY 1644 AAGGCTCCTGGGTTCCAGCACC 1703
DB 1805 AAGGCTCCAGGCTTCCAAACACC 1864
QY 1704 TCAGCTGCTCTTACTAATGAC 1763

Db	1965	TCAGCAGTCCCTTACTAATCGAGCCTTCTCTGTTTAAATCTTGATGTGATCACTACATAAAT	1926
Qy	1764	AACAGCAAGGCTGTCGAGAAGCTATGTGCTTCTCTAATAGGATCCAAACCTAGGTCCGCAA	1823
Db	1925	AATAGCAAGCCTTTAAGGGAAGCTATGTGTTTCTGTATGGATCCAAACCTTGGGAAGCAA	1984
Qy	1824	GTCTGTTATGTGGAGTTTCCCAACAAGGTTTGATGGGATTGATAGGAATGATCGATGCA	1883
Db	1985	GTTTGTATAGTTCAGTTCCTCTCAAGAGTTTGATGTGATATGATCGGAATGATCGATATGCC	2044
Qy	1884	AACAGGAACACTGTCTTTTTTTCATATTACTGTGAGGCGCTTGACGGCATTCARAGACCA	1943
Db	2045	AACAGAAATACCTTTTTTTCATATAAATCTTGAGGGTCTGGATGGCATTCAGAGGCGG	2104
Qy	1944	GTTTATGTGGAACTGGTTGTGTTTCAACAGAACAGCTATCTATGTTATGATGAGCCGCCA	2003
Db	2105	GTTTATGTGGTACGGATGTGTTTTCATAGACAGCATTTGATGATATGAACACCCGCC	2164
Qy	2004	ATTAAAGCGAAG-----AAGCGAGGTTTCTTGGCATCACATATGTGGGG---CAAGAAG	2054
Db	2165	CTTAAGCCCAAGCATAGAAAAACAGGATACTATCTTCATTGTGGGAGGTTCTCGTAAG	2224
Qy	2055	AAGCRAAGCAAGTCAAGAAGAGGAGCTCAGATAGAAAAAGTCCGAACAAGCATGTGCGAC	2114
Db	2225	AAGAGTCAAAATCAAGTAAAAAGGATCAGAACAGAAATACTGGCAAGCATTTGTGAT	2284
Qy	2115	AGTCTCTGTTCCAGTATTCAACTCGAAGACATPAGAGGAGGTTGTGAAGTGTGCTGGGTTT	2174
Db	2285	TCTACTGTACCAGTATTCATCTAGAGGATATAGAAGAGGTTGTGAAGGGGCTGGGTTT	2344
Qy	2175	GATGATGAGAAATCAGTTCTCATGTCTCAATAGCTTTAGAGAAGAGATTTGGCCAGTCA	2234
Db	2345	GATGATGAAAAATCAATTACTGATGTACAAATAGCCCTGGAGAAAGGTTTGGCCAGTCT	2404
Qy	2235	GCAGCATTTGTGCCCTCCACTCTGATGGAATATGTGTGTTCTCAGTCTCCTCACATCCA	2294
Db	2405	GCGTGTTTTGTGGCCCTACACTTATGGAATAATGGTGTGCCCTCAGTCTGCTACGCC	2464
Qy	2295	GAATCTCTTTTGAAGAAGCTATCAGTCATCAAGTCTGTGCTATGAGGACAAGTCTGAA	2354
Db	2465	GAACACTCTTTAAGGAGCAATTCATGTTATCAGCTGTGGATAGGAAGACAAACAGAC	2524
Qy	2355	TGGGGAACCTGAGATTGGTTGGATCTATGGATCTGCACAGAAGATATTTCTTACTGGATTC	2414
Db	2525	TGGGGAAGTGAGATTGSGTGGATCATGTTCTGTTACAGAAGATATTTCTTACTGGATTC	2584
Qy	2415	AAGATGACGCAAGAGGCTGGGCTCAGTCTATTCGATGCCCAAGCGGCCAGCTTTCAG	2474
Db	2585	AAATGCATGCCGCTGGTTGGGGGTCAATTTACTGCATGCCAAACGCGCCATTTAAA	2644
Qy	2475	GGATCTGCCCCCATCAATCTTTCAGATCGCTGACCAAGTGTCCGTTGGGCTCGCTCGGT	2534
Db	2645	GGTTCGTCTCATCAATCTTTCAGATCGCTGGAACCAAGTGTTCGATGGGCTCTTGGT	2704
Qy	2535	TCTGTTGAAATCTTTTTCACGCGGATTTGCCCTTTATGGTATGGTACGGAGGCGGCTC	2594
Db	2705	TCTGTGAAATCTTTTTCAGTAGGCAATGGCCCTATCTCGTATGGTATAGTGGTAGGCTC	2764
Qy	2595	AAGTTCCTGGAGAGATTCCGTTACATCAACACCACCATTTACCACATTAACCTCTCTCCGG	2654
Db	2765	AAGTGCCTTGAAGCATTCGCATATGTTAACACAACCATCTACCCAGTCACTGCTATTTCCT	2824
Qy	2655	CTTCTAGTCTATTGTATATTGCTGCTATCTGCTGCTCACTGGAAAGTTTCATCATGCCA	2714
Db	2825	CTTCTCATGTACGTACACTGCCACAGTCTGCTGCTCACTACACAAGTTCTATTATTTCCA	2884
Qy	2715	GAGATTAGCAACTTGGCCAGTATCTGGTTTCATTGGGCTCTTCCCTTCAATTTTCCGCACT	2774
Db	2885	CAGATTAGTAACCTTGCAGTATATGGTTATATCCCTCTTCTCCTCACTTTTGCACCT	2944
Qy	2775	GGTATCCTTGAGATGAGGTGGATGGTGTGTGGCATTTACGAGTGGTGGAGGAATGAACAG	2834
Db	2945	GGTATTTTAAAAATGAAGTGGAAATGGTGTGGGAAATGACCAATGGTGGAGAAATGAACAG	3004

Qy	2835	TTCTGGGTCATTGGAGGATATCTGCGACATCTGTTTGGCGGTCFTTTCAGG	2894
Db	3005	TTTTGGGTTATCGGTGGTGTGCGGCACATCTTTTGTGTTTTCCTCAAGGCGCTTCTCAA	3064
Qy	2895	GTGCTTGGCGGTATCGACACCAACTTCACTGTCACTCAAGGCTAATGACGAGAAAGGC	2954
Db	3065	GTTCCTTGTGGCATTTGACACTTAATTCACGGTCACTTCAAAGCATCAGACGAAGATGGG	3124
Qy	2955	GACTTTGCTGAGCTCTACATGTTCAAGTGGACGACGGTCTTCTCATCCCTCCGACGACCAT	3014
Db	3125	GATTTGCTGAATTTACATGTTCAATGGACACCCCTTCTCATCCCAACCAACCATCTC	3184
Qy	3015	TTGAFATTAACATGGTTGGTGTGCTGTGGACACTCTTACGCCATCAACAGATGGTTAC	3074
Db	3185	CTCATATAAACTTTGGTGGGGTGTGTTGCAGGAATATCGTATGTCATCAATAGTGGCTAC	3244
Qy	3075	CAATCATGGGGCGCGCTCTTTTGGGAAGCTCTCTTGCCTTCTGGTGCATGTTCACTTA	3134
Db	3245	CATCATGGGGTCCCTCTTTGGTAAGTAATCTTGGCTTCTGGGTGATCATCCATCTC	3304
Qy	3135	TACCCATTCTTCAAGGCTTATGGCAGGCAAAACCGCACCGAGCATTTGTCATCGTC	3194
Db	3305	TACCCCTTCTTGAAGGTTTATGGAGCGCCAAACCGTACACCTACCATTTGTTGTA	3364
Qy	3195	TGGGCTGTCTCCGCTTCTATCTTCTCTTGTGCTGTGGTGTGCTGTGATCCATTCCT	3254
Db	3365	TGGTCAATCTGCTGTGCTCCTATTTCTCTTTGTTGGTGGGATGATCCCTTCACA	3424
Qy	3255	ACCGTCTCGCTGGGCCAAATATCCAAACCGTGGCATCAACTGCTAGGAAAG	3307
Db	3425	ACCAGATTACCGGCGGATGTGAACAGTGTGAATAAACTGTTGAGAGAG	3477
RESULT	6		
LOCUS	AF027174	3682 bp	linear
DEFINITION	Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B)		
ACCESSION	AF027174		
VERSION	AF027174.1		
KEYWORDS	GI:2827142		
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (bases 1 to 3682) Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J. and Williamson, R.E.		
TITLE	Molecular analysis of cellulose biosynthesis in Arabidopsis		
JOURNAL	Science 279 (5351), 717-720 (1998)		
MEDLINE	98111412		
PUBMED	9445479		
REFERENCE	2 (bases 1 to 3682)		
AUTHORS	Arioli, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-SEP-1997) Plant Science Centre, Australian National University, Acton, Canberra, ACT 200, Australia		
FEATURES	Location/Qualifiers		
source	1..3682 /organism="Arabidopsis thaliana" /cultivar="Columbia" /db_xref="taxon:3702"		
gene	1..3682 /gene="Ath-B"		
CDS	256..3453 /gene="Ath-B" /note="RSWI-like" /codon_start=1 /product="cellulose synthase catalytic subunit" /protein_id="AAC39336.1"		

QY 1824 GTCTGTTATGTCAGTTCACCAAGGTTTGATGGATTGATAGGAATGATCGATATGCA 1883
DB 1969 GTTTGTTATGTTTCAGTTCACCAAGATTGATGGATTGATAGGAATGATAGATATGCT 2028
QY 1884 AACAGGAACACTGCTCTTTTGTATATACTTGAAGGGCTTGACGGCATTCACAGGACCA 1943
DB 2029 AATCGTAATACCGTGTCTTTGATATTAATCTGAGAGGTTTAGATGGGATTCAAGGACCT 2088
QY 1944 GTTTATGTGGAACTCGTTGTGTTTCAACAGAACAGCTATCTATGGTTTATGAGCCCCA 2003
DB 2089 GTATATGTCGGAACCTGGATGCTTTTCAACAGAACAGCATATACGGTTTATGAACCTCA 2148
QY 2004 ATTAAGG- - - - - CGRAGAGCCAGGTTTCTTGGCATCTATGTGGGGGCAAGRAG 2057
DB 2149 ATAAAGTAAACACAAAGAACGAGTCTTTATCTAAAGCTCTGTGGTGGATCAAGAAAG 2208
QY 2058 GCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAAAGTCGAACAAGCATGTGGACAGT 2117
DB 2209 AAGAATTCAAAAGCTAAGAAAGAGTCGGACAAAAAATCAGGCAGGCATCTGACTCA 2268
QY 2118 TCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGGTGTGAAGGTGCTGGGTTTGAT 2177
DB 2269 ACTGTTCTGTATTCATCAACTCGATGATAGAGAGGGAGTTGAAGGTGCTGGTTTGAT 2328
QY 2178 GATGAGAAATCAGTCTCTATGTCCTCAAAATGAGCTTAGAAGAAGATTTGSCCAGTCAGCA 2237
DB 2329 GATGAAGAGCGCTTTATGTCGCAATGAGCCTGGAGAAGGATTTGGACAGTCTGT 2388
QY 2238 GCATTTGTTGCCCTCCACTCTGATGGAATATGTTGTTTCTCCAGTCCCTCCACTCCAGAA 2297
DB 2389 GTTTTGTGTTCTTACCCTAATGGAATATGTTGTTTCTCTTTCAGCAACTCCAGAA 2448
QY 2298 TCTCTTTTGAAGAAGCTATCCATGTCATGAAGTTTGGCTATGAGGACAAAGTCTGAATGG 2357
DB 2449 AACCTTCTCAAGAGGCTATCCATGTCATTAGTTGTGTTATGAGGATAAGTCAGATTGG 2508
QY 2358 GGAATGAGATTGTTGGATCTATGATCTGTCCACAGAAGATTTCTTACTGGATTCAAG 2417
DB 2509 GGAATGAGATTGATGGAATCTGTCACAGAAGATTTCTTACTGGATTCAAG 2417
QY 2418 ATGCACGCAAGAGGCTGGGCTTCAGTCTATTGTCATGCCCAAGGCCAGCTTTCAAGGGA 2477
DB 2569 ATGCATGCCGCTGGATGGGATCCATTACTGTCATGCCTAAGCTTCCAGCTTTCAAGGCT 2628
QY 2478 TCTGCCCCCATCAATCTTTCAGATCTGCTGAACCAAGTGTGGGTGGGCTCTCGGTTCT 2537
DB 2629 TCTGCTCCTATCAATCTTTTCAGATCTGCTGAACCAAGTGTGAGGTGGGCTTTAGGTTCA 2688
QY 2538 GTTGAATCTTTTCAGCCGGCATTCGCCCTTATGTTGCTACGGAGGCGCCCTCAAG 2597
DB 2689 GTTGAGATTCTTTCAGTCGGCATTCCTATATGTTATGTTTACAAATGGGAGGCTAATA 2748
QY 2598 TTCCGTGGAGATTCCGTTTACATCAACACACACATTTACCACATTAACCTCTCTCCCGCTT 2657
DB 2749 TTTCTTTGAGAGGTTTCGGTATGTGAACACACCATCTACCCCTATCACCCTCCACTTCTCTT 2808
QY 2658 CTAGTCTATTGTTATGTCGTCTATCTGCTCTCACTGGAAGTTTCATATGCCAGAG 2717
DB 2809 CTGATGTTTGTACATTGTAGCCGTTGTCTCTTCCACCAACCAAGTTTATTTATCTCTCAG 2868
QY 2718 ATTAGCAACTTGCCCAAGTATCTGTTTCAATTCGCTCTTCTCTTCAATTTTTCGCCACTGGT 2777
DB 2869 ATTAGTAACATTGCAAGTATATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2928
QY 2778 ATCCCTTGAGATGAGGTGGAGTGTGTTGGCATTGACAGTGTGGAGGAATGAACAGTTTC 2837
DB 2929 ATACTAGAAATGAGGTGGAGTGGCGTAGGCAATAGACGAATGGTGGAGAAACAGCAGCTTT 2988
QY 2838 TGSGTCATTGGAGTATCTCTGCACATCTGTTTGGCCGTCTTTCAGGGTCTTCTGAAGGTG 2897
DB 2989 TGGGTCATTGGTGAGATCCCGCTCATTTATTTCGCTGTGTTTCAAGGTATCTCTCAAGTCA 3048

QY 2898 CTTGCCGCTATCGACACCAACTTCACCTGCTCACCTCAAAAGGCTAATGACGAAGAGCGAC 2957
DB 3049 CTTGCCGCTATGACACAACACTTCACATTCACCTCAAAAGCTTCAGATGAAGACGGAGAC 3108
QY 2958 TTTGCTGAGCTCTACATGTTTCAAGTGGACGAGCTTCTTCATCCCTCCGACGACCACTTTG 3017
DB 3109 TTTGCTGAGCTCTACTTGTTCAAATGGACAACACTTCTGATTCGCCCAACGACGCTGCTC 3168
QY 3018 ATCATTAACATGTTGGTGTGCTTGGGACCTCTTACGCCCATCAACAGTGGTGTACCAA 3077
DB 3169 ATTGTAACCTTAGTGGAGTTGTTGACAGGAGTCTCTTATGCTATCAACAGTGGATACAA 3228
QY 3078 TCATGGGGCCGCTCTTTTGGGAAGCTCTTTCGCTTCTGCGGTGATGTTGTTCACTTATAC 3137
DB 3229 TCATGGGACCACTCTTTGTAAGTGTCTTTCGCTTCTGGGAGTGTGTTCACTTGTAC 3288
QY 3138 CCATTCTCAAGGCTTTATGGGAGGCAAAACCGCACACCGAGCATTTGTCATCTGCTGG 3197
DB 3289 CCTTCTCCCAAGGTTTATGGTGGCGACAGAACCGGACTCTACCATTTGTTGCTGCTGG 3348
QY 3198 GCTGCTCCTCGCTTCTATCTTCTCTTCTGCTGCTGGGTTGCTGTTGATCCATTCACACC 3257
DB 3349 TCTGTTCTTCTGCTTCTATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3408
QY 3258 GCTCTCGCTGCGCCAAATATCAAAACCTGTGGCATCAACTGCTAGGAA 3305
DB 3409 CGAGTCACTGGCCGGGACATTCGGAATCTGGAATCAACTGTTGAGAA 3456

RESULT 7
AX030946
LOCUS AX030946 3614 bp DNA linear PAT 20-SEP-2000
DEFINITION Sequence 9 from Patent WO9800549.
ACCESSION AX030946
VERSION AX030946.1 GI:10278349
KEYWORDS thale cress.
SOURCE ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 3614)
Williamson,R.E., Peng,L., Arioli,A. and Betzner,A.S.
Manipulation of cellulose and/or beta -1,4-glucan
Patent: WO 9800549-A 9 08-JAN-1998;
WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI
ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;
COMWU SCIENT IND RES ORG (AU)
Location/Qualifiers
1. 3614
/organism="Arabidopsis thaliana"
/strain="COLUMBIA"
/db_xref="taxon:3702"
/clone="ATH-B"
217. .3414
/note="unnamed protein product"
/protein_id="CAC09627.1"
/db_xref="gi:10278350"
/translation="MESEGETAGPKMKNIYVPTQICSDNVGKTVGDREVAICISF
PVCRCYERKDNQSCPOCKTRKRLKSPAIIPDKDEGLADEGTVEFNYPQEK
ISERMLGWLHTRGKEEMGPQYDKVEYSHNHLRLTSRQDTSGEFSFSAASPR
TAGKRLPYSDVNQSPNRRIYDPVGLGNVAKERVDPGKWKOEKNTGVPYSTOASER
GGVIDASTDILADEALLNDEARQLLSRKYSIPSSRLNPYRMVLTMLRLVLCFLHYR
ITNPVNAFLWLVSVICEIWFALSWILDQFPKFPVNRREYLDRLALRDREGEPSQ
LAAVDIFSVTDPLKEPLTANTVLSIADYDPVKVSCVSDGGAAMI.SFSLAET
SEFARKWPKCKYSEIAPRAEPYFAAKIDYLDKQVTSFVKDRKRAKREYEEFKIRI
NALYSKALCKPEGVWQDGPWPNNGTDPHGMIOVFLQNGGLDAGEMLRLVVV
SREKRPGFQHHKAGAMNALYRSVLTNGPFIILNDCDHVINNSKALREAMCFMDP
NLKQVCYVQFPORFDIGDKNDRYANRNTVFYFLLNRLGLDGIQGPVYVGTGCVNRTA
LYTEPPIKVKHAPSLLSLCGGSRKSKAKKESDKSKRSRSTSTVPVFNLDLIE
EGVREGAGFDDEKALLMSQLEKFRFGQSAVFVASTLMENGVPPSPATPEFLKPAIH
ISCYEDKSDMGMEIGWYISVTEILTGFKMHARGWRSIYCMKLPFAFGSAPINLS

DRNLQVLRNALGSVEILPFRSHCDIWIYNGRLKFLERFAYVNTIYPIPSILPMYCT
LIAVCLTFNQFIPOISNIAISWFLSFIATDGLMRWSGVGDENWRNQFVVI
GGVSAHLFAVSVIAGIDNTFTVTSKASDEGDFAELYFKWTTLLIPPTLLI
NLVGVVAGVSVAINSGYSGWPLFKLFFAFWVIVHLYPFLKGLMRQNRRTIIVV
WSVLLASIESLLWRIIDPFTSRVTGPDILLECGINC.
BASE COUNT 963 a 717 c 863 g 1071 t

ORIGIN	Query Match	Best Local Similarity	71.4%	Pred. No. 0;	Mismatches 869;	Indels 72;	Gaps 8;
QY	30	TCGGTGGCGTGGATCGCTCTCGCGGCCATGGACGGCGACGGCGGACGCCCTGGAATCC	89				
Db	190	TCGGTGTGGGAACAATAAGTGAACAAATGGAATCCGAAGGAGAAACCGCGGGAACCG	249				
QY	90	GGGAGGACGGCGCGGGGACGTGTGCCAGATCTGCCCGACGGCCTGGGCACCACTTG	149				
Db	250	ATGAAGAACATTTGTCGCGAGACTTGCAGATCTGTAGTACAAATGTTGGCAAGACTGTT	309				
QY	150	GACGGGACGCTCTTCAACCGCTGCGAGCTGTGCCGTCTCCCGTCTGCCGCCCTGCTAC	209				
Db	310	GATGGAGATCGTTTGTGGCTTGTGATATTTGTTCATCCAGTTTGTCCGCCCTTGTCTAC	369				
QY	210	GACACGAGCGCAAGGAGGACCCAGCGCTGCTCCAGTGCAGAGACCAAGTACAAGCGC	269				
Db	370	GAGTATGAGAGGAAGATGGAAATCAATCTGTCTCAGTGCAGTCAAAACCAAGAGG	429				
QY	270	CACAGAGGACCGACGATCCCGGGGAGGAGGCGACACACTGATGCCGATGATGGT	329				
Db	430	CTCAAAGGTAGTCTCTATTCCTGTGTGATAAAGACGAGGATGGCTTAGCTGATGAAGT	489				
QY	330	A---GTGACTTCAACTACCCCTGCTGGCACTGAGACCAAGAGCAGAGATTGCTGAC	386				
Db	490	ACTGTTGAGTCAACTACCCCT-----CAGAAGGAGAAATTTCAAG	531				
QY	387	AGATGGCAGCTGGCGCATGACACCGGGGCGAGTGGCAATGTGGCCACCCCAAGTAT	446				
Db	532	CGATGCTTGGTGGCATCTTACTCTGTGGGAAGGAGAGAAATGGGGAAACCCAGTAT	591				
QY	447	GACATGGCGAGATCGGCTCTCCAAAGTATGACAGTGGAGATCCCTAGGGATACGTC	506				
Db	592	GATAAGAGG-----TCTCTCAATCATCTTCCCTGCTCAGGACGACACAAG--	640				
QY	507	CCTTCAGTCAACCAACCGCAGATGTGAGGAGAAATCCCTGGAGCTTCGCCCTGATCATC	566				
Db	641	-----ATACTTCAGGAGAGTTTCTGCTGCCCTCACTGACGCCCTC	681				
QY	567	ATGATGTCCCTACGGGGAACATCAGACAGCTGCTCGGTTTCCCT---ATGTGAATCAT	623				
Db	682	TCTGTATCTTACTATATCGGTGGGAAAGCGCCTTCCCTATTCATCAGATGTCAATCAA	741				
QY	624	TCACCAAAATCCGTCAGAGGAGTCTCCGGCAGTATTTGGGAATGTTGCCGTGGAAAGAGA	683				
Db	742	TCACCAATAGAAGGATGTGGATCTGTGGACTCGGGAATGATGCTTGGAAAGAGAGA	801				
QY	684	GTTGATGGCTGGAATAATGAAGCAGGACAAAGGTGGCATGCCATGACTAATGGCAACGC	743				
Db	802	GTTGATGGCTGGAATAATGAAGCAGGACAAAGTACTGTGCTT-----TGTCAACAG	852				
QY	744	ATTGCTCCCTCTGAAGGTGGGAGCTACTGACATGATGATCTACTGAATACACATG	803				
Db	853	CAGGCTGCTTCTGAAAGAGGTGGAG---TAGATATTGATGCCAGCACAGATATCCTAGCA	909				
QY	804	GAAGACGCTTTACTGAATGATGAATCGCCAGCCTCTATCTAGAAAAGTCCCATGCT	863				
Db	910	GATGAGGCTCTGCTGAATGACGAAGCGGCGAGCTTCTGTCAAGGAAAGTTTCAATTCCT	969				
QY	864	TCTCCAAAATAAATCCCTACGAATGGTCAATGTTCTCGGGTGGTGTGTTCTTAAGCATC	923				
Db	970	TCATCAGCGATCAATCCTTACAGATGGTTATTATGCTCGGGCTGTTATCCTTCTCTC	1029				
QY	924	TTCCCTGCACTACCGTCTCAAAATCCTGTGGTGAATGCAATACCCACTGGGCTTTTATCT	983				

Db	1030	TTCTTGCATTACCGTATAACA	1089	AGTCCAAATGCCTTTGCTCTATGCTGCTCT
QY	984	GTATATGTGAGATTTGGTTG	1043	ATCTGGATACTGGATCAGTTCCTCCCAAGTGGTTT
Db	1090	GTGATGTGAGATCTGGTTG	1149	ATCTGGATTTGGATCAGTTCCTCCCAAGTGGTTT
QY	1044	CCATCAACCGCGGAGACTAC	1103	TAGACTGGCTTTAAGGTATGACCCGAGAAGGTGAA
Db	1150	CTGTGAACCGTGAACCTAC	1209	CAGGCTTGCCTTTAAGATATGATCGTGAAGGTGAG
QY	1104	CGGTCTCAGTTGGCTGCTGTT	1163	ATTGTGACGTACAGTCGACCCCTTGAAGGAGGCA
Db	1210	CCATCAGATGATGCTGTTG	1269	TTTCTGAGTACTGTTGACCCCTTGAAGGAGGCA
QY	1164	CTATCTCCTACTGCCAACACT	1223	ATCCATCTCTCTGTTGATTATCCCGTGGACAG
Db	1270	CCCTTGTGACAGCAACAG	1329	CTCTATTCTGCTGTTGACTATCCCAAGTTCAGAG
QY	1224	GTCTCTTGTCTATGATCTGAT	1283	AGCTTCAATGCTGACTTTTGTGACGATTTGGCTGAG
Db	1330	GTGCTGTTATGTTCTGATG	1389	GTCTGCTATGTTATCATTTGATCACTTGCAGAA
QY	1284	ACTTCAGAGTTTGTAGAAAT	1343	ACCATTTTGTGAAGAAATATAGCATAGAGCCCTCGT
Db	1390	ACATCAGATTTGCTCGTAAAT	1449	ACCATTTTGTGAAGAAATATAGCATAGAGCCCTCGT
QY	1344	GCCTCCGAGTTTACTTTTGGC	1403	ATTTGATTACCTGAAAGACAAAGTCCAGCCCTTCA
Db	1450	GCACCAAGATTTGGTACTTTGCT	1509	ATAGATTACTTTGAAGGATTAAGTTTCAGACATCA
QY	1404	TTTGTAAAGACCGCCCGGCA	1463	SAGAAATATCAAGAAATTTAAATCAGGATAAAT
Db	1510	TTTGTCAAAGATCGTAGAGCT	1569	SAGGAAATATCAGGAATTTAAATCCGAAATCAAT
QY	1464	GCCTAGTTTCTTAAGCATTTGA	1523	SCCCGAGGAAGATGGATCATGCAAGATGGCACA
Db	1570	GCCTTTTCCAAAGCCCTTAA	1629	CCCTGAAGAGGGTGGGTTATGCAAGATGGCACA
QY	1524	CCATGCCAGGAAACATACCA	1583	TCATCCTGGATGATTGAGGTTTTCCTTGGTGCAC
Db	1630	CCGTGGCCTGGAATAATACAG	1689	SCATCCAGGAATGATCCAGGCTTCTTTAGGCA
QY	1584	AGTGGTGGCTTGTACTGAGG	1643	TGAGTCCCCCTTTAGTTTATGTCCTGCTGAA
Db	1690	AATGGTGGACTTGTGTCAGAG	1749	TGAGTCCCCGCTTGGTATATGTTTCTCGAGAA
QY	1644	AAGGCTCCTGGGTTCCAGCAC	1703	SAAGGCTGGTGGCATGAATGCCCTTGTGTTGTC
Db	1750	AAGCGACCGAGATTCACGACC	1809	SAAGGCTGGTGGTATGAATGCACTGGTGAGAGTT
QY	1704	TCAGCTGCTCTTACTAATGAG	1763	ATGTTGAAATCTTGTGATGATCACTACATCAAC
Db	1810	TCAGCAGTCTTACCACATGGA	1869	ATCTTGAATCTTGTGATGATCACTAATCAAT
QY	1764	AACAGCAAGCTCTCCGAGAG	1823	ATGCTTCTTAATGATCCAAACCTAGGTCCGCA
Db	1870	AACAGCAAGCTTAAAGAGAG	1929	ATGCTTCTGATGAGCCCAACCTCGGGAAGCAA
QY	1824	GTCGTTATGTGAGTTTCCAC	1883	ATTTGATGGGATGATAGGAATGATGATGCA
Db	1930	GTTTGTATGTTTCCAGTCCCA	1989	ATTTGATGGTATCGATAAGAACGATAGATATGCT
QY	1884	AACAGCAACTGCTCTTTTTC	1943	TACTTGAAGGGCTTTCAGCGCATTCAGGACCA
Db	1990	AATCGTAATACCGTGTCTTTC	2049	TACTTGAAGGGTTCAGTGGGATTCAGGACCT
QY	1944	GTTTATGTGGAACTGTTGTC	2003	TAACAGAACAGTATCTATGTTTATGAGCCCCCA
Db	2050	GATATGTGCGAACTGGATGTC	2109	TAACAGAACAGCATTATACGGTATGAACTCCA
QY	2004	ATTAAGG-----CGAAGAGG	2057	TTCTTGGCATCACTATGTGGGGGCAAGAGAA

Db 2110 ATAAAGTAAACACACAAAGCCAAAGTCTTTATCTAAGCTCTGTGGTGGATCAAGAAAG 2169
QY 2058 GCAAGCAAGTCAAGAAAGGAGCTCAGATAGAAAGAAAGTGAACAAAGCTGTGGACAGT 2117
Db 2170 AAGAATTCCAAAGCTAAGAAAGAGTCGGAACAAAGAAATCAGGCAGGCATCTGACTCA 2229
QY 2118 TCTGTTCCAGATTAATCTCGAAGACATAGAGGAGGTGTGAAGGTGCTGGGTTCAT 2177
Db 2230 ACTGTTCTGTATTAACCTCGATGACATAGAAAGAGGAGTGAAGGTGCTGGTTCAT 2289
QY 2178 GATGAGAAATCATGTTCTCATGTTCAATGAGCTTAGAGAGAGATTTGGCCAGTCAGCA 2237
Db 2290 GATGAAAGGCCCTCTTAATGTCGAATGAGCTTGGAGAGCGATTTGGACAGTCTGCT 2349
QY 2238 GCAATGTTGCTCCACTCTGAATGATATGATGTTGTTCTCTGATGCTCTCCACTCCAGAA 2297
Db 2350 GTTTTGTGTTCTACCTAATGAAATGTTGTTGTTCTCTCTCTGACCACTCCAGAA 2409
QY 2298 TCTCTTTGAAGAGAGCTATCCATCTCATAGTTCTGCTATGAGGACAGTCTGAATGG 2357
Db 2410 AACTTCTCAAAGAGGCTATCCATGTCATAGTTGTTGTTATGAGGATAGTCAGATGG 2469
QY 2358 GGAATCGATGATGTTGGTATCTATGATCTGTACAGAGATATTTCTACTGATTCAG 2417
Db 2470 GGAATGGAGATGGATGATCTATGTTCTGTGACAGAGATATTTCTGACTGGTTCAAA 2529
QY 2418 ATGACGCAAGAGGCTGGGTTTCACTATATTCATGCTCCCAAGCCGCGGCTTTCAAGGA 2477
Db 2530 ATGCATGCCGCTGGATGGGATTCATTTACTGATGCTTAAGCTTCCAGCTTTCAAGGT 2589
QY 2478 TCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTCGGTGGCTCTCGTTCT 2537
Db 2590 TCTGCTCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTGAGTGGCTTTAGTTCA 2649
QY 2538 GTTGAATCTTTTTCAGCCGCAATGCCCCCTTATGATGCTACGAGGCGGCTCAAG 2597
Db 2650 GTTGAGATCTCTTCAGTCGGCATGTTCTATATGATGTTTACAACTGGGAGGATAAA 2709
QY 2598 TTTCTGGAGAGATTCGCTTACATCAACACCAACCATTTACCCACTAACCTCTCTCCGCTT 2657
Db 2710 TTTCTTGAGAGGTTTTCGATGTGACACCAACCATCTACCCCTATCACCTCCATCTCTT 2769
QY 2658 CTAGTCTATTGATATGCTGCTATCTCTGCTCACTGGAAGTTTATCATGTCAGAG 2717
Db 2770 CTATGATTTGATGTTGCTAGCGCTTCTCTCTTCCACCAACAGTTATTTATCTCTCAG 2829
QY 2718 ATTAGCAACTTGGCCAGTATCTGTTTCATGTCGCTCTTCCCTTCAATTTTCCGCACTGGT 2777
Db 2830 ATTAGTAACATTGCAAGTATATGTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
QY 2778 ATCCTTGAGATGAGTGGAGTGGTGTGGCATTGACGAGTGTGGAGGAATGAACAGTTC 2837
Db 2890 ATACTAGAATGAGTGGAGTGGCTGAGCATAGACGAATGTTGGAGAAACGAGAGTTT 2949
QY 2838 TGGCTCATTTGGAGTATCTCTGCATCTGTTTGGCGTCTTTCAGGCTTCTTGAAGTGT 2897
Db 2950 TGGCTCATTTGGAGTATCCGCTCATTTATTCGCTGCTTTCAGGATCTCTCAAGTCT 3009
QY 2898 CTTGCGCGTATGACACCAACTTCACTGTCACCTCAAGAGCTTAATGACGAAGAGCGAC 2957
Db 3010 CTTGCGCGTATGACACCAACTTCAAGTTTACCTCAAAAGCTTCAGATGAAGACGAGAC 3069
QY 2958 TTTGCTGAGCTTACATGTTCAAGTGGACGAGCTCTCTATCCCTCCGAGCACTTTTG 3017
Db 3070 TTTGCTGAGCTTACATGTTCAATGGACAACACTTCTGATTCGCGCAACAGCTGCTC 3129
QY 3018 ATCATTAATGATGTTGGTGTGCTGTCACCTCTTACGCACTCAACAGTGGTTACCAA 3077
Db 3130 ATTGTAACACTTAGTGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCAA 3189
QY 3078 TCATGGGGCCGCTCTTTTGGAGAGCTCTTTTGGCTCTGCTGGTGAATGTTCACTATAC 3137
Db 3190 TCATGGGGCACTCTTTTGGTAAAGTGTCTTTTGGCTCTTCTGGTGTATGTTCTCACTGTAC 3249

QY 3138 CCATTCTCTAAGGCTCTTATGGGCGAGGCAAAACCGCACCGGAGATTGTCTATCGTCTGG 3197
Db 3250 CTTTCTCTCAAGGCTTTGATGGTTCGACAGACCGGACTCTTACCATTTGTGGTCTGG 3309
QY 3198 GCTGCTCTCTCGCTTCTATCTCTCTCTGCTGCTGGTCTGCTGTATGATTCATTACATTAC 3257
Db 3310 TCTGTTCTCTGCTTCTATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3369
QY 3258 CGTCTGCTGCGCCAAATATCAAAACCTGTGGCATCAACTGCTAGGAA 3305
Db 3370 CGAGTCACTGGCCGGACATTCGTGAATGTGGAATCAACTGTTGAGAA 3417

RESULT 8
AX030960
LOCUS AX030960
DEFINITION Sequence 9 from Patent WO9800549.
ACCESSION AX030960
VERSION AX030960.1 GI:10278363
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (Bases 1 to 3614)
Williamson, R.E., Peng, L., Arioli, A. and Betzner, A.S.
Manipulation of cellulose and/or beta -1,4-glucan
Patent: WO 9800549-A 08-JAN-1998;
WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI
ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;
COMWU SCIENT IND RES ORG (AU)

FEATURES
source
1. 3614
/organism="Arabidopsis thaliana"
/strain="COLUMBIA"
/db_xref="taxon:3702"
/clone="ATH-B"
217..3414
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09633.1"
/db_xref="GI:10278364"

CDS

translation="MESEGTACKPMKNTVPOTCIQSDNVGKTVGDRFVACDICSF
PVCRCPEYERKDNQSCPOCKTRYKELGSPAIPGDKDDEGLADEGTFEYFNYPQKEH
ISERMLGHLTRGGEEMPEPQDKVEYSHNHLRLSRQDTSFEFSAAPERLSVST
IAGGRLPYSSDVNQPNRRIVDPGLGNVAMKERVDMKMKOEKNTGVPSTQAAER
ITNPVNAFALWLVSVICETWALSWILQFPKFPVNRETYLDRLALRYDREGPSQ
LAADVDFSVTDLPKPPVTANTVLSILADVDPVKVSYSDSDGAAMLSPESLAET
SEFARKWPECKKYSIEPRAPENYFAKIDYLDKVQTSFKVDRRAKREYEFKIRI
NALYSKALKCPESGWMDGTPWPNNTGDHPGMIQVFLGQGLDAGNELPRLVYV
SRKRPGFHKKAGAMNALVRVSAVLNTGPFILNLDCHYINNSKALREAMCFLMDP
NLGKQCVYQFPQDFGIDKNDRYANRNTFFDINLGLDIOGQVYVGTGCVFNRTA
LYGVEPIKVKHKKPSLLKLCGSRKSKAKSKKSKGSRHTDSTVYGVNLDIE
EGVEGAGDEKALLMSLEKRFQGVSAFVASTLMENGVPSPATPENFLKALH
ISCGYEDKSDWMEIGWYGVSTEDILTGPKHARGWSLYCMKLPAPKSGAPNLS
DRNLQVLRWALGSVEILFNRHCPIWYNGYRLKFLERFAVNTTIPISILMNTCT
LLAVGFTNQFIIPQISNIASIMFLSLFATGILEMWSGVDEWMNRNQFWI
GGVSAHLFAVGILKVIAGIDNTFTSKADEGDFAELYFKWTTLLIPIPTLLI
VNLVGAVGVSVAINSQYQSWGPLFGKLPFAFWIVHLYPLFLKGLMGRNRTPTIVV
WSVLASIFSLNVRIDPFTSRVTPGDILECGINC"

BASE COUNT 963 a 717 c 863 g 1071 t
ORIGIN

Query Match 46.2%; Score 1673.6; DB 8; Length 3614;
Best Local Similarity 71.4%; Pred. No. 0;
Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;

QY 30 TCGGTGCGGTTGGATTCCTCTGCCGCCCATGACGCGGACGCGGAGCTCTGAAGTCC 89
Db 190 TCGGTGTTGGAAGCAACTAAGTGACAAATCGAATCCGAAGGAGAAACCGCGGAAAGCCG 249

QY 90 GGGAGCGCGGGGACGTGTGCCAGATCTGCGCCGACGCGCTGGGCACCACGTTG 149
DB 250 ATGAAGAACATGTTCCGACACTGCGACATCTGTAGTGACAATGTGGCAAGACTGTT 309
QY 150 GACGCGACGCTCTTACCGCCTGGGAGCTGTGCGGCTTCCGGTCTGCGGCCCTGCTAC 209
DB 310 GATGGAGATCGTTTGTGGCTGTGATATTGTTCAATCCGAGTTTGTGCGGCCCTGCTAC 369
QY 210 GAGCAGCGACGAGGAGGACCGACGCTGCTCCAGTCCAGACGACGATACAAGCGC 269
DB 370 GAGTATGAGAGGAAGATGGAATCAATCTTGTCTCAGTGCAAAACGAGATACAAGAGG 429
QY 270 CACAGAGGCCCGCAGCGATCCGGGGGAGGAGGAGGACGACACTGATGCCGATGATGGT 329
DB 430 CTCGAAGGTAGTCTGCTATTCCTTGGTGATTAAGACAGGAGTGGCTTAGCTGATGAAGGT 489
QY 330 A---GTGACTTCAACTACCTGCATCTGGCACTGAGGACCAAGACAGAGATGCTGAC 386
DB 490 ACTGTTGAGTTCAACTACCT-----CAGAAGGAGAAAATTTCAAG 531
QY 387 AGGATGCCACTGGCGCATCAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTAT 446
DB 532 CGGATGCTGTTGGCATCTTACTCGTGGGAAGGAGGAGAAATGGGGGAACCCCAAGTAT 591
QY 447 GACAGTGGCAGATCGGCCTCTCCAAGTATGACAGTGGAGAGATCCCTAGGGGATACGTC 506
DB 592 GATAAAGAGG-----TCTCTCACAATCATCTCTCGCTCACGAGCAGACAAG-- 640
QY 507 CCTCAGTCAACACAGCCAGATGTGAGGAGAAATCCCTGGAGCTTCGCCGTGATCAAC 566
DB 641 -----ATACTTCAGGAGGTTTCTGCTGCCTCACCTGAAACGGCTC 681
QY 567 ATGATGCTCCCTACGGGAACATCAGCAGAGCTGCTCGCTTTCCT---ATGTCGAATCAT 623
DB 682 TCTGATCTTCTACTATCGCTGGGGAAAGCGCCTTCCCTATTATCAGATGCTCAATCAA 741
QY 624 TCACCAAAATCGTCAAGGAGTTCTCCGGCAGTATTTGGGAATGTGGCTTGGGAAAGAGAGA 683
DB 742 TCACCAAAATAGAAGATTTGTGGATCTCTTGGACTCGGGAATGTAGCTTGGGAAGAGAGA 801
QY 684 GTTATGCTCGAAATGAACGACGACAGAGGTCGATTCCTCATGACTAATGCGCAGAC 743
DB 802 GTTGATGCTGGAATGAACGACAGAGAGAAATACTGGTCC-----TGTGAGCAGC 852
QY 744 ATTGCTCCCTCTGAAGTCTGGGACGCTACTGACATCGATGATGATCTACTGATACAACATG 803
DB 853 CAGGCTGCTTCTGAAAGAGGTGGAG---TAGATATTGATGCCAGACAGATATCCTAGCA 909
QY 804 GAAGACGCTTACTCAATGATGAACCTGCCAGCCTCTACTAGAAAAGTCCCATTTGCT 863
DB 910 GATGAGGCTGTGTAATGACGAGCGGAGGCGCTTCTGTAAGGAAAGTTTCAATTTCT 969
QY 864 TCCTCCAAAATAATCCCTACAGAATGTCTATGTTCTCGGTTGGTTTCTTAAGCATC 923
DB 970 TCATCAGGATCAATCTCTACAGAATGTTATATGCTCGGCTGTATCTCTTGTCTC 1029
QY 924 TTCCTGCACTACCGTCTCAGAAATCTCTGCGGTAAATGCAATCCCACTGTGGCTTTATCT 983
DB 1030 TTCCTGCAATACCGTATACCAACCCAGTGCCAAATGCTTGTCTATGGCTGTCTCT 1089
QY 984 GTTATATGAGATTTGGTTTGTCTTATCTCGGATCTGGAATCAGTTCCCGAAGTGGTTT 1043
DB 1090 GTGATATGAGATCTGTTTGGCTTATCTCTGGATTTGGATCAGTTTCCCAAGTGGTTT 1149
QY 1044 CCAATCAACGGGAGACCTACTCTGATAGACTGGCTTTAAGGTATGACCGAGAAAGTGAA 1103
DB 1150 CCGTGAACCGTGAACCTACTCTGACAGGCTTGCCTTAAGATATGATCGTGAAGGTGAG 1209
QY 1104 CCGTCTCAGTTGGCTGTGTTGACATATTTCTCAGTACAGTCGACCCCTTCAAGAGGCCA 1163
DB 1210 CCATCAGACTAGCTGCTGTTGACATTTTCTGTGAGTACTGTTGACCCCTTGAAGAGGCCA 1269

QY 1164 CCTATCGTCACTGCCAACACTG 1164
DB 1270 CCCCCTGTGACAGCCACACAG 1270
QY 1224 GTCTCTGCTATGTATCTGATG 1224
DB 1330 GTGCTCTGTTATGTTTCTGATG 1330
QY 1284 ACTTCAGAGTTTCTAGGAAT 1284
DB 1390 ACATCAGAGTTGCTGTAAT 1390
QY 1344 GCTCCCGAGTTTACTTTTGC 1344
DB 1450 GCACCAGATGTTACTTTGCTG 1450
QY 1404 TTTGTTAAAGACCCCGGGCC 1404
DB 1510 TTTGTCAAGATGTTAGAGCT 1510
QY 1464 GCCTAGTTTCTAAGCATTT 1464
DB 1570 GCACTTTTCCAAAGCCCTTA 1570
QY 1524 CCATGGCCAGGAACRATACCA 1524
DB 1630 CCGTGGCTGGAATAATACAG 1630
QY 1584 AGTGGTGGCTTGTATCTGAGG 1584
DB 1690 AATGGTGGACTTGTATGAGAG 1690
QY 1644 AAGCGTCTCGGTTCCAGCAC 1644
DB 1750 AAGCGACAGATTCACAGCAC 1750
QY 1704 TCAGCTGCTCTTACTAATGAGC 1704
DB 1810 TCAGCAGTTTCTCCAAATGAG 1810
QY 1764 AACAGCAAGGCTGTCCGAGAG 1764
DB 1870 AACAGCAAGCCTTAAGAGAG 1870
QY 1824 GTCTGTTATGTGAGTTTCCAC 1824
DB 1930 GTTGTGTTATGTTCAAGTTCC 1930
QY 1884 AACAGGAACACTGTCTTTTGT 1884
DB 1990 AATCGTAAATCCGTTGTTCT 1990
QY 1944 GTTATATGGGAACTGGTTGTC 1944
DB 2050 GTATATCTCGAACTGGATGTG 2050
QY 2004 ATTAAGG-----CGAAGAGC 2004
DB 2110 ATAAAGTAAACACAAAGAGC 2110
QY 2058 GCAGCAAGTCAAGAAAGAGAG 2058
DB 2170 AAGAATTCAAAGCTAAGAAAG 2170
QY 2118 TCTGTTCCAGTATTCATCTCG 2118
DB 2230 ACTGTTCTCTATTTCAACCTCG 2230
QY 2178 GATGAGAAATCAGTTCATGTG 2178
DB 2290 GATGAAAGGGGCTCTTAATGT 2290
QY 2238 GCATTTGTTGCCCTCCACTCTGA 2238

ATCCATTCTTGTCTGTTGATTATCCGTTGGACAAG 1223
CTCTATTCTGCTGTTGCTGACTACCCAGTTGACAAG 1329
AGCTTCAATGCTGACTTTTGGCGCATTTGGCTGAG 1283
TGTGCTATGTTATCATTTGAATCACTTGCAGAA 1389
ACCATTTGTGAAGAGTATGACATTTGAACCCAGA 1343
ACCATTTTGCAGAAATATAGCATAGAGCCCTCGT 1449
AATTGATTACCTGAAGACAAAGTCCAGCCCTCA 1403
AATAGATTACTTGAAGGATTAAGTTTCAGACATCA 1509
SAGAGAAATGAAGAAATTTAAATCAGAGATAAT 1463
SAGGAATATGAGGAATTTAAATCCGAATCAAT 1569
SCCGAGGAAGGATGGATCATGCAAGATGGCACA 1523
CCCTGAAGAGGGTGGTTATGCAAGATGGCACA 1629
CATCTCTGGAATGATTCAGGTTTCTTGGTGCAC 1583
CATCCAGGAATGATCCAGGTTCTTTTAGGCA 1689
GAGCTCCCCGTTTGTATGTTGCTCTCGTGAA 1643
GAGCTCCCGGTTTGGTATATGTTCTTCGAGAA 1749
AAGGCTGGTGCATGAATGCCCTTGTTCGTGTC 1703
AAGSCTGGTCTATGAATGCACCTGGTGAGATT 1809
ATGTTGAATCTTGAATGTTGATCATCATCAAC 1763
ATCTTGAATCTTGAATGTTGATCATCATCAAT 1869
TGCTTCTTAATGATCAACACCTTAGTCCGCA 1823
TGCTTCTGATGACCCCAACCTTCGGAAGCA 1929
TTTGATGGGATGATAGGAATGATCATGATGCA 1883
TTTGATGGTATCGATAGAAGCATAGATATGCT 1989
AACTTGAGGGCTTGTGACGGCATTCAGGACCA 1943
AACTTGAGAGGTTTGTGACGGCATTCAGGACCT 2049
AACAGAACAGCTATCTATGTTTATGACCCCA 2003
AACAGAACAGCATTATACGGTTATGAACCTCA 2109
TTCTTGGCATCACTATCTGCTGGGCAAGAGAAG 2057
CTTTTATCTAAGCTCTGTGGTGGATCAAGAAAG 2169
GATAAGAAAAGTTCGAACAGCATGTGGACAGT 2117
GACAAAAGAAATCAGGCGAGCATCTGACTCA 2229
ATAGAGGAGGCTTGAAGGCTGCTGGTTTGTAT 2177
ATAGAAGAGGAGGTTGAAGGCTGCTGGTTTGTAT 2289
ATCAGCTTACAGAAAGAGATTTGGCCAGTCAAGCA 2237
ATGAGCTTGAGAAAGCGATTTGGACAGTCTGCT 2349
ATAGGTTGGTGTCTTCCAGTCTCTCCACTCCAGAA 2297

Db 2350 GTTTTGTGCTCTTACCCCTAATGGAATGFGGTTCTCCTCTCAGCAACTCAGAA 2409
QY 2398 TCTCTTTTGAAGAAGCTATCCATGTCATAAGTTGTGGCTATGAGGACAAGTCTGAATGG 2357
Db 2410 AACTTCTCAAGAGGCTATCCATGTCATTAGTTGTGGTTATGAGGATAAGTCAGATTGG 2469
QY 2358 GGAACCTCAGATTGTTGGATCTATGATCTGTGCACAGAGATATCTTACTGGATCAAG 2417
Db 2470 GGAATGAGATTGGATGGATCTATGTTCTGTGCACAGAGATATCTGACTGGGTCAAA 2529
QY 2418 ATGCACCAAGAGGCTGGCTTTCAGTCTATTGTCATGCCCAAGGCCAGCTTTCAGGGA 2477
Db 2530 ATGCATGCCGCTGGATGGGATCCATTACTGTCATGCCCTAAGCTTCCAGCTTTCAGGCT 2589
QY 2478 TCTGCCCCCATCAATCTTTCAGATCGTCTGACCAAGTGTGGGGTGGCTCTCGGTTCT 2537
Db 2590 TCTGCTCTCAATCTTTCAGATCGTCTGACCAAGTGTGGGGTGGCTTTCAGGTTCA 2649
QY 2538 GTTGAATCTTTCAGCGGCTTTCAGCTTATGCTATGCTTACGGCTACGGAGGCGCTCAAG 2597
Db 2650 GTTGAGATTCTTTCAGTGGGATGCTCTATGATGTTTACATGGGAGGCTTAA 2709
QY 2598 TTCTGAGAGATTCGCTTACATCAACACACCATTTTACCCACTAACCTCTCTCCCGCTT 2657
Db 2710 TTTCTTGAGAGGTTTGGGTATGTGAACACACCATCTACCTTACCTCCATCTCTCTT 2769
QY 2658 CTAGTCTATTGATGCTGCTATCTGCTCTCTACCTGGAAGTTTATCATGCGAG 2717
Db 2770 CTGATGATTGATGATGCTGCTGCTCTCTTACCAACACCATTTATTTCTCTCAG 2829
QY 2718 ATTAGCAACTTGGCGGCTATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2777
Db 2830 ATTAGTAACATTGCAAGTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889
QY 2778 ATCTCTGAGATGAGTGGATGCTTGTGGCATTTGACAGTGTGGGAGGAATGAACAGTTC 2837
Db 2890 ATACTGAATGAGTGGGAGTGGGATGAGCATAGAGCAATGTTGGGAGAAACGAGCATTT 2949
QY 2838 TGGTCTATTCGAGTATCTCTGACATCTGTTTGGCGTCTTCTTCTTCTTCTTCTTCTT 2897
Db 2950 TGGTCTATTCGAGTATGCTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3009
QY 2898 CTGCGCGGTATGACACCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2957
Db 3010 CTGCGCGGTATGACACCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3069
QY 2958 TTTGCTGAGCTTACATGTTCAAGTGGACAGCGCTTCTCATCCCTCCGAGCAGCATTTG 3017
Db 3070 TTTGCTGAGCTTACTTGTTCAAATGGACAACTTCTGATTCGCGCAACGAGCGTCTC 3129
QY 3018 ATCAATTAACATGTTGTTGCTGTTGCTGGCACCTCTACGCCATCAACAGTGTACCAA 3077
Db 3130 ATGTAAACATTAGTGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGTATACCAA 3189
QY 3078 TCATGGGGGCGCTCTTTGGGAAGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3137
Db 3190 TCATGGGGGCGCTCTTTGGGAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3249
QY 3138 CCATTCCTCAAGGCTTATGGGCGACGCAAAACCGCACCGGAGTGTGTCATGCTCTGG 3197
Db 3250 CCTTCTCTCAAGGTTGATGGGTCGACAGAACCGGAGCTCTACCATTTGTTGTTGTTCTGG 3309
QY 3198 GCTGCT 3257
Db 3310 TCTGTTCT 3369
QY 3258 CGTCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACTGTGTAGAA 3305
Db 3370 CGAGTCACTGGCCCGGACATTCTTGGAAATGTGGAATCACTGTGTAGAA 3417

RESULT 9

AF200526
LOCUS AF200526 3725 bp mRNA linear PLN 31-AUG-2000
DEFINITION Zea mays cellulose synthase-2 (Cesa-2) mRNA, complete cds.
ACCESSION AF200526
VERSION AF200526.1 GI:9622875
KEYWORDS
SOURCE
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3725)
AUTHORS Holland, N., Holland, D., Helentjaris, T., Dhugga, K.S., Xoonostlie-Cazares B. and Delmer, D.P.
TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene family
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)
MEDLINE 20398328
PUBMED 10938350
REFERENCE 2 (bases 1 to 3725)
AUTHORS Dhugga, K.S. and Helentjaris, T.G.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA
FEATURES
source
1. .3725
/organism="Zea mays"
/db_xref="taxon:4577"
/note="Identified and isolated from the maize genome database cDNA collection of Pioneer Hi-Bred International, Inc."
gene
1. .3725
/gene="Cesa-2"
179..3403
/gene="Cesa-2"
/note="Cesa-2"
/codon_start=1
/product="cellulose synthase-2"
/protein_id="AAF89962.1"
/db_xref="GI:9622875"
/translation="MAANKWAGSHNRNEFVMIRHGDAPVPKAPTKSANGVCQICGDVGVATGDVFNACNECAFFPCYCYERKEGNOCCQCKTRYKROKSGSRVHGDDEEDVDLDNEFNKYGNGKPEQWQDDADLSSARHDPHRIPLRTSGQISGEIPASDPDRHSIRSTSYSDVPVPIRVDPDKLNSYGLNSVDKRWKRVKQDKNMLQNTKYPEARGMEGTGNGEDMOMVDARLPLSRIVPISSNQLNLYRIVILRLIILCFEFOYRI SHPVNAYGLMLVSVICVWFALSKLLDQPKWPIINRETVLDRLARLDREGEPSQLAPIDVFVSTVDPLKEPLLTANTVLSLAVDYPDKVSCYSDSDGSAMLTFSLSSETAEFAKWPFCRKHNIETPEAFEFYFAQKIDYLDKIQIPSVFKERRAMKREYEFKIRINALVAKAQVPEEGTMDAGTAMPNPNRPHGMIOVFLHSGGLDTDGNEPLRVTVYSREKRFQHHKAGAMNALLIRSVALTNGAYLLAVDCHDFNSSKLREAMCPMDPALGRKTCYVQFPORFDGIDLHRYANRNIVFFDINMGJLDGQGPVYVGTCCFNQALYGYDPLVLEADLEPNIVVKSCCRKRKNKSYMDSQSRIMKRTSSAPFNEMEDIEGIEGDEDSVLSMSORKLEKRGQSPFIASFMTQGGIPPTNPASLLKEAIIHVSIGEDKTEWKEIGWIVGSVTEDLITGFMHARGWSIYCMPPRPFKGSAPINLSRLNQLRWALGSVEILSRHCPWIYNGYNGRLKLERLAYINTIVYPTTSVPIATCYLPAICLLTNKFIPEISNAGMFFILFASIPATGILRLWSGIEDMW RNEQFWVIGGTSAHFLFAVFOGLKVLGIDTNETVTSKSDEGDAEILYVFKWTSILLIPPTVLVINLGMVAGISVAINSGYSGMGLFKLFFSIWILHYLPFLKGLMGRQNRTFTIVINSILLASIPSLWVKIDFISFTQKAAALGCGVNC"
BASE COUNT 927 a 807 c 950 g 1041 t
ORIGIN
Query Match 38.8%; Score 1408.4; DB 8; Length 3725;
Best Local Similarity 66.8%; Pred. No. 3.7e-267;
Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;
QY 53 CCGGCCCATGGACGCGACCGCCCTTCAAGTCCGCGAGGACGCGGCGGGGACGT 112
Db 238 CCGCACGACGCGACGCGCCCTGTCGCCGCTAAGCCACGAAGAGTGCGAATGGCAGGT 297
QY 113 GTGCCGATCTGCGCGACGCGCTGGGACACCATGTTGACGCGGACGCTCTTCACCGCCTG 172
Db 298 CTGCCAGATTTGTGCGACACTGTTGGCGTTTTCAGCCACTGTTGTTGCTGCTG 357

QY 173 CGACGCTGCGCTTCCCGTCTGCGCCCTGCTAGAGCAGCGGCAAGGCGCAC 232
DB 358 CAATGAGTGTGCCCTTCCCTGTCTGCGCCCTTCTGTATAGTAGAGCGCAAGGAAGGAA 417
QY 233 CCAGGCGTGCCTCCAGTGGCAAGACCAAGTACAAAGCGCCAGAGGAGCCAGCGATCCG 292
DB 418 CCAATGCTGCCCTCAGTGCAGAACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTCA 477
QY 293 CGGGGAGGAAGCGGACGACACTGATGCCGATGATGGTAGTGTCTCAACTACCCCTGCATC 352
DB 478 TGGTGATGATGAGGAGGAAGATTGTGATGACCTGGCAATGAATCAACTATATAGCAAGG 537
QY 353 TGGCACTGAGGACCAAGACAGAGAGATTGCTGACAGGATCGCGAGCTGGGCGCATGACAC 412
DB 538 CAATGGGAAGGGCCAGAGTGGCAGCTT-----CAAGGAGAT 574
QY 413 CGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACATGGCGAGATCGGCCTCTCCAA 472
DB 575 GACGCTGATGCTGCTTCATCTGCTGCCATGACCCACACCATCGGATTCACGCCCTTACA 634
QY 473 GTATGACAGTGGAGATCCCTAGGGGATACGTCCTTCAGTCACCAACAGCCAGATGTC 532
DB 635 AGTGGACA-----ACAGATATC 651
QY 533 AGGAGAAATCCCTGGAGCTTCCGCTGATCATCATGATGTCCTCCCTACGGGGAACATCAG 592
DB 652 TGGAGAGATCCCTGATGATCCCTGACCGTCACTTATCCCGCAGTCCCAACATCGAGCTA 711
QY 593 CAGACGTGCTCGCTTCCCTATGTGAATCATTCACCAATPCCCTCAAGGGAGTT---CTC 649
DB 712 TGTGTATCCAAGGGTCCAGTTCCTCTGAGGATTGTGGACCCCTCGAAGGACTTGAATTC 771
QY 650 CGGCAGTATTGGGAATGTTGCCCTGGAAAGAGAGAGTGTGCTGGAAATGAAGCAGGA 709
DB 772 CTATGGCTTAATAGTGTGACTGGAAGAAAGAGTTGAGAGCTGGAGGTTAAACAGGA 831
QY 710 CAAGGCTGGGATPCCCATGACTAATGGGACAGCATTTGCTCCCTCTGAAGCTCGGCAGC 769
DB 832 CAAAAATATGTTGCAAGTGACTAA-----TAAATATCCAGAGGC 870
QY 770 TACTGACATCGATCTACTGAATACAAACATGGAGAGCGCTTTACTGAATGATGAAC 829
DB 871 TAGAGAGACATGGAGGCGATGGCTCAAAATGGAGAGATATGCAAAATGGTTGATGATGC 930
QY 830 TCGCCAGCCTCTATCTAGAAAAGTCCCAATGCTTCTCCCAAAATAAATCCCTACAGAAAT 889
DB 931 ACGCCCTACCTTTGAGCGGCAATTGCGCAATTTCTCAACACGCTCAACCTTTACCGGAT 990
QY 890 GGTCAATGTTGCGGTTGGTGTCTAAGCATCTTCCTGCACTACCGTCTCACAATCC 949
DB 991 AGTAATCATCTCCGCTTATCATCTCTGCTCTTCTTCCCAATATCGTATCAGTCATCC 1050
QY 950 TGTGCTAATGATACCCACTGCGCTTTATCTGTATATGAGATTGGTTTGGCTTT 1009
DB 1051 AGTGCCTAATGCTATGGAATTTGGCTAGTATCTGTATCTGTGAGGCTGFGTTTGGCTTT 1110
QY 1010 ATCCTGGATACTGGATCAGTTCGCCGAAGTGGTTTCCAAATCAACGGGAGACCTTACCTTGA 1069
DB 1111 GTCCTGCTCTAGATCAGTTCCCAAAATGGTATCCAAATCAACCGTGAGACATATCTCGA 1170
QY 1070 TAGACTGGCTTTAAGTATGACCGAAGGTGAACCGTCTCAGTTGGCTGCTGTGACAT 1129
DB 1171 CAGGCTTGCATTTAGGTATGATAGAGAGGAGAGCCATCAGACTGGCTGCCATTTGATGT 1230
QY 1130 ATTTGTCACTACAGTACCCCTTGAAGGAGCCACCTATCGTCACCTGCCCAACACTGTGCT 1189
DB 1231 CTTTGTGACGTACAGTGGATCCATTTGAAGGAACCTCCACTGATCACAGCCAACTGTTTT 1290
QY 1190 ATCCATCTTGTGTTGATTATCCCTGGACAAAGGCTCTTGTATCTATCTGATGACGG 1249
DB 1291 GTCCATCTTGCTGTGATTAACCTCTTGACAAAGTGTATGTTCTGATGATGG 1350

QY 1250 AGCTTCAATGCTGACTTTTGAC 1309
DB 1351 CTGAGTATGCTGACTTTTGAC 1410
QY 1310 ACCATTGTGAAGAAGTATGAC 1369
DB 1411 TCCTTTTGTGAAGGCACAAAT 1470
QY 1370 AATTGATTACCTGAAGACAAAT 1429
DB 1471 AATAGATTACTTGAAGCAAAAT 1530
QY 1430 GAGAGAAATGAAGAATTTAAAT 1489
DB 1531 GAGAGATTATGAAGAATTTCAAT 1590
QY 1490 CCCGAGGAAGGATGGATCATG 1549
DB 1591 GCCTGAAGAGGGTGGACCATG 1650
QY 1550 TCATCCTGGAATGATTCAGGTT 1609
DB 1651 CCATCCTGGCATGATTCAGGTT 1710
QY 1610 TGAGCTCCCCCGTTTGTGTTTAT 1669
DB 1711 TGAATTACCAAGCTCTTGTCTAT 1770
QY 1670 GAAGGCTGGTGCCATGNAATCC 1729
DB 1771 GAAGGCTGGTGCAATGAATGCA 1830
QY 1730 CATGTTGAATCTTGTGTTGAT 1789
DB 1831 TCTTCTCAATGTGATTTGATG 1890
QY 1790 GTCTTCTTAATGGATCCAAAG 1849
DB 1891 GTGCTTTCATGGATCCAGCT 1950
QY 1850 GTTGTGATGGATTCATAGGAAT 1909
DB 1951 ATTTGATGGCATTGACTTGGAC 2010
QY 1910 TAACTTTGAGGGGCTTGACGGC 1969
DB 2011 CAACATGAAGGCTAGATGCG 2070
QY 1970 CAACAGAACGCTATCTATGCT 2026
DB 2071 CAATAGGCGAGCTTGTGATGGA 2130
QY 2027 CTTTGGCATCACTATGTGGGGC 2086
DB 2131 CATTTGTTTAAAGCTGCTGTG 2190
QY 2087 TAAGAAAAAGTTCGAACACGAT 2146
DB 2191 TCAAGCCGATTTATGAAGAGA 2250
QY 2147 AGAGGAGGTTGTTGAAGGTCCT 2206
DB 2251 CGAGGAGGTTTGAAGGTT--- 2304
QY 2207 GAGCTTAGACAAGAGATTTGGC 2266
DB 2305 GAAATGGAGAAACGCTTTGGT 2364
QY 2267 TGGTGTGTTTCCCTCAGTCTCC 2326
DB 2365 AGGTGGCATACCACCTTCAACA 2424
QY 2327 AAGTTGTGGCTATGAGGACAA 2386

QY 293 CGGGAGGAAGGACGACACTGATCCCGATGATGGTAGTGACTCAACTACCCCTGCATC 352
DB 482 TGGTGATGAGGATGAGGAAGATTGATACCTAGACAAATGAATCAACTAC-----533
QY 353 TGGCACTGAGGACGACAGACAGAGATGCTGACAGGATGCGCAGCTGCGCGATGAACAC 412
DB 534 -----AAGCAAGC 542
QY 413 CGGGGGCAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGCTGGCTCTCCAA 472
DB 543 AGTGGGAAGGCCAGAGTGGCAACTGCAAG---GAGATGATGCTGATCTGCTCACT 599
QY 473 GTATGACAGTGGAGAGATCCCTAGGGGATACGTCCTTCAGTCAACACAGCCAGATGTC 532
DB 600 GCTGCCATGAGGCACATCATCGGATTCACACGCTGACACGGGTCAA----CAGATATC 655
QY 533 AGGAGAAATCCCTGGAGCTTCGGCTGATCATCACATGATGTCCTTACCGGGGAACATCAG 592
DB 656 TGGAGATGCTGATGCTTCCTCCCTGACCGTCACTTATCCGAGTCCCAACATCGAGCTA 715
QY 593 CAGAGTGCCTGCTTCCCTATGTAATCATTCACCAAAATCCGTCAAGGGAGTTC---TC 649
DB 716 TGTGATCCCAAGGCTCCAGTTCCTGTGAGGATGTGGACCTCGAAGGACTTGAATTC 775
QY 650 CGGCAGTATTGGGAATGTCCTGGAAAGAGAGAGTGTGATGGCTGGAAATGAACAGCA 709
DB 776 CTATGGCTTAATAGTGTGACTGGAAGGAAGAGTGTGAGCTGAGGGTTAAACAGGA 835
QY 710 CAAGGGTGGCATCCCATGACTAATGGGACAGCATGTGCTCCCTGTAAGGTGCGGGCAGC 769
DB 836 CAAAAATATGATCAAGTCAAGTACTAATAATATCCAGAGGCTGAGGAGGAGATGGAGGG 895
QY 770 TACTGACATCGATGATCTACTCAATACACATGGAAGAGCGTTTACTCAATGATGAAC 829
DB 896 GACTGGCTCAATGGA-----GAANATATGCAAAATGGTGTGATGATGC 937
QY 830 TCGCCAGCCTCTATCTAGAAAATCCCATTTGCTTCCCTCAAAATAAATCCCTACAGAT 889
DB 938 ACGGCTACCTTTGAGCGGTATCGTGCCAATTTCTCAAAACAGCTCAACCTTTACCGGGT 997
QY 890 GGTCAATGTTGGGTGTTCTTCTAAGCATCTTCCCTGCACTACCGTCTCACAAATCC 949
DB 998 AGTATCATCTTCGCTTATATCATCTGCTTCTTCCAGTATCGGTGAGTATCATCC 1057
QY 950 TGTGCGTAATGATACCATCGCTGCTTTTATCTGTATATGATGAGATTTGTTGCTTT 1009
DB 1058 AGTGGGTGATGCTATGGAATGCTAGTATCTGTTATCTGCGAGGTCTGTTTGCCTT 1117
QY 1010 ATCTGTGATGATGATGCTCCGGAAGTGGTTTCCAAATCAACCGGGAGACCTACCTTGA 1069
DB 1118 GTCTTGGCTTCTAGATCAGTTCCCAAAATGGTATCCAAATCAACCGTGAGACATATCTTGA 1177
QY 1070 TAGACTGGCTTAAAGGTATCAGCGAGGTGAACCGTCTCAGTTGGCTGCTGTTGACAT 1129
DB 1178 CAGGCTTGCAATGAGGTATGATAGAGGAGGAGACCATCACAGCTGGCTCCCATGATGT 1237
QY 1130 ATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
DB 1238 CTTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1297
QY 1190 ATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249
DB 1298 GTCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1357
QY 1250 AGCTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
DB 1358 TTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1417
QY 1310 ACCATTTGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
DB 1418 TCCCTTTGTGAAGACACAATATGAACCAAGAGCTCCAGATTTTACTTTGCTCAAAA 1477

QY 1370 AATTGATTACCTGAAGACAAAT 1429
DB 1478 AATAGATTACTGAAGGACAAAT 1537
QY 1430 GAGAGAATATGAAGAATTTAAAT 1489
DB 1538 GAGGAGATGAAGAATTCAAAT 1597
QY 1490 CCCCAGGAAGGATGGATCATG 1549
DB 1598 GCCTGAAGAGGGTGGACCATG 1657
QY 1550 TCATCCTGGGAATGATTCAGGTT 1609
DB 1658 CCATCCTGGGATGATTCAGGTT 1717
QY 1610 TGAGTCCCCCGTTTATGTTTAT 1669
DB 1718 TGAGTTTACCAGCTCTTGTCTAT 1777
QY 1670 GAAGGCTGGTCCCATGAATGCC 1729
DB 1778 GAAGGCTGGTCCCATGAATGCC 1837
QY 1730 CATGTTGAATCTTGAATGATGAT 1789
DB 1838 TCTTCTCAATGTGGATGCGAC 1897
QY 1790 GTGCTTCTTAAATGATCAAAAC 1849
DB 1898 GTGCTTCTTAAATGATCAAAAC 1957
QY 1850 GTTGTGATGGATGATGATGATGAT 1909
DB 1958 ATTTGATGGATGATGATGATGATGAT 2017
QY 1910 TAACTTGAAGGCTTGAAGGCTTGA 1969
DB 2018 CAACATGAAGGCTTGAAGGCTTGA 2077
QY 1970 CAACAGAACAGCTATCATGTTG 2026
DB 2078 CAATGACAGCTTGTGATGATGAT 2137
QY 2027 CTGTCATCATCTATGTTGGGGCT 2086
DB 2138 CATTTGTTTAAAGAGCTGCTGCT 2197
QY 2087 TAAGAAAAGTCCGAACAGCATG 2146
DB 2198 TCAAGCCGTTATGATGAGAGAT 2257
QY 2147 AGAGGAGGCTTGAAGGCTGCTG 2206
DB 2258 CGAAGAGGTTTCAAGST--- 2311
QY 2207 GAGCTTAGAAGAGATTTGGGCT 2266
DB 2312 GAAATTTGGAAGAGCTTTGGTGT 2371
QY 2267 TGGTGGTGTCTTCCAGTCCCTCC 2326
DB 2372 AGGTGGCATACCATCTTCAACAA 2431
QY 2327 AAGTTGGCTTATGAGGACAGT 2386
DB 2432 CAGTTGTGGATATGAGGACAAAT 2491
QY 2387 TGTCAAGAGATATCTTACTG 2446
DB 2492 AGTAACGAGGATATCTGACTG 2551
QY 2447 TTGCATGCCCAAGGCCCGCTT 2506

Db	2552	CTGCATGCCACACACACCTTGTTCACAGGGTCTGCACCAATCAATCTTCGGATCGCT	2611
Qy	2507	GAACCAAGTGTGCGGTGGCTCTCGGTCTGTTGAAATCTTTTTCAGCCGGCATGGCC	2566
Db	2612	TAATCAGGTGCTCCGTTGGCTCTGGTCACTGAGTGAATCTCTAGTAGACATGTCC	2671
Qy	2567	CTTATGGTATGCTACGAGGCGCTCAAGTCTCGAGAGATTCGCTTACATCAAC	2626
Db	2672	TATCTGGTATGTTTACAATGACCATGAAGCTTTTGAGAGGCTGCTTACATCAAC	2731
Qy	2627	CACCATTTACCACATACTCTCCCGCTTCTAGTCTATTGTATATTCGCTGCTATCG	2686
Db	2732	TATTGTATATCAATCAATCCATTCGCTTATTGGCTATTGTGCTTCCGCTATCG	2791
Qy	2687	TCGTCTCACTGAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTCA	2746
Db	2792	CTCTCTTACCAATAAATTATCATCTCTGAGATTAGCAATATATGCTGGGATGTTCTCA	2851
Qy	2747	TGCGCTCTTCTCTTCAATTTTGGCCTCTGATCTCTGAGATGAGTGGAGTGGTGTGG	2806
Db	2852	TCTTCTTTTGGCTTCCATTTTGGCCTCTGATATGAGCTTAGATGGAGTGGTGTGG	2911
Qy	2807	CATTGACGAGTGGTGGAGATGAACAGTCTCGGCTCATTTGGAGGTATCTCTCCACATCT	2866
Db	2912	CAATTTTGGCTTCCAGGCTCTCTGAAAGTCTTGGCTTATTTGGTGGCACCCTCTGCCATCT	2971
Qy	2867	GTTTGGCGTCTTTCAGGGTCTTCTGAAGGTCTTTCGCGGTATCGACACCAACTTCACTGT	2926
Db	2972	CTTTCGAGTGTTCAGGCTCTCTGAAAGTCTTGGCTTGGATGATACCAACTTCACTGT	3031
Qy	2927	CACCTCAAGGCTAATGAGGAAGAGGCGACTTTGCTGAGCTCTACATGTTCAAGTGGAC	2986
Db	3032	TACCTCAAGGCTATGATGAGGATGGCGACTTTGCTGAGCTATGTTGTTCAAGTGGAC	3091
Qy	2987	GAGCTTCTTCTCCCTCCGACGACCATTTTGCATTAACATGTTGGTGTCTGCTGTGG	3046
Db	3092	CAGTTTGTCTCATCTCCGACCACTGTTTGTTCATTAACCTGCTGGAATGTTGGCAGS	3151
Qy	3047	CACCTCTACGCCATCAACAGTGGTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTT	3106
Db	3152	AATTTCTATGCAATACAGATGGCTACCAATCTGGGGTCCGCTCTTTGGAAGCTGT	3211
Qy	3107	CTTTGCTCTCTGGGTGATTTGCTTATACCCATTCCTCAAGGGTCTTATGGCGAGGA	3166
Db	3212	CTTCTCGATCTGGGTGATCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAAGGA	3271
Qy	3167	AAACCCACACCGAGATGTCATGCTGGGCTGCTCTGCTGCTTCTATCTTCTCTT	3226
Db	3272	GAACCCGACACCAACAATCGTATGCTGTGCTCATCTCTTGTGCTCATCTCTTCTCTT	3331
Qy	3227	GCTGTGGGTCTGTTGATCCATTCATCACTACCC	3258
Db	3332	GCTGTGGGTGAAGATCGATCTTTCATCTCCC	3363
RESULT	11		
LOCUS	AX030942		
DEFINITION	Sequence 5 from Patent WO9800549.	DNA	linear
ACCESSION	AX030942		
VERSION	AX030942.1	GI:10278345	
KEYWORDS			
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 3603)		
AUTHORS	Williamson, R.E., Peng, L., Arioli, A. and Betzner, A.S.		
TITLE	Manipulation of cellulose and/or beta -1,4-glucan		
JOURNAL	Patent: WO 9800549-A 5 08-JAN-1998; WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI		

ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;
COMMW SCIEN IND RES ORG (AU)
Location/Qualifiers
1..3603
/organism="Arabidopsis thaliana"
/strain="COLUMBIA"
/db_xref="taxon:3702"
/clone="RSW1 CDNA"
1..3246
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC09625.1"
/db_xref="GI:10278346"
/translation="MEASGLVAGSVRRNELVRIHRESGGTGPLKNNMGQICQICGD
DVGLAETGDFVACNECAFPCPCYEYERKDTGCCPOCKTFRHRHSRPRVGGDED
EDVDVDIENEYVAGANKARHQBHEFSRSSHSOPILPLTHGHTVSGEIRPDPT
QSVRTTSGPLGSDRNAISSPYIDPRQPVPRIVDFSKDLNSGLNVDNKKERVEGW
LKOENKMLQMTGKYHKGEGIEGTSGBELQADDRLPMSRVVTPFSSRLTPTRV
VILRLIILCFLOYRTHPVKNAYPLWLTSTVCEIWFASFLLDDQFPKWPINRETY
LDLRAIRYDRGEPQSLVPDVFVYFVDPLKEPLVANTVLSILSDYVDKQSVK
SDGSAMLTFESLSEAFKFKVFCFKFNIEPRAFEPYFAQIDYLDKQIPSEVK
ERRAMKREYEEFKVRIKALVAKOIKPEEGTWMODGTPWPNNTRHGMIQVFLGHS
GLDLDGNELPRLIYVSRKRPQHHKAGAMNALIRVSALTNAGYLLNVDQDHYF
NNSKAIKAMCFMDPAIGKCCYQEPQEDGIDLHDYANRNIYFEDINMKGLDGI
QCPVTYGGCCFNQALYGPVTEEDLEPNIIYKSCSRKKGSKKSKYNERRG
INRSDNAPLFNMEDIDEGFYDDERSILMSQSRSEKREGQSPVIAATFMEQGGIP
PTNPATLLKEAIIHIVSCGYEDKTEWKEIGWYGVSTEDILTGFKMHARGWISYCN
PPRPATKGSAPINSLRNQVLRWALGSEILLRSHCPWIWYGHRLERLRIAYINT
IYVPTTSLIAYCILPAFLITDRFIPEISNVASITWILFLEISITAVTGILRWMSG
VSIEDWRNEOFWYIGTSAHLFAVFOGLLKLADIDTNFTVTSKATDDEGDFAELYI
FKWTLALIPPTVLLNVLIGIVAGSVAVNSYOSMGPLFGKLFALWIAHLYLPLK
GLLGRQNTPTPIVWMSVLLASIESLLWLRINPVDANPNANFNKGGVF"
BASE COUNT 950 a 745 c 852 g 1056 t
ORIGIN
Query Match 37.0%; Score 1343.4; DB 6; Length 3603;
Best Local Similarity 66.2%; Pred. No. 2.4e-254;
Matches 2098; Conservative 0; Mismatches 991; Indels 80; Gaps 8;
Qy 92 GAGGCACGGGCGCGGAGCGTGTGCCAGATCTGCCGCCAGCGCTGGCGACACAGTTGGA 151
Db 93 GAAGAATPATGAATGGCCAGATATGTCAGATCTGTGGTATGATGTGGACTCGCTGAAC 152
Qy 152 CGCGAGCTTTCACCGCTGGACGCTCTGCCGCTTCCCGCTCTCCGCCCTCTCTACGA 211
Db 153 TGGAGATGCTTTGTGCGGTGAATGAATGTGCCCTTCCCTGTGTGCGCTTGTATGA 212
Qy 212 GCACGACGCAAGAGGGACCCAGCCCTGCCCTCCAGTGCAGACCAAGTCAAGCGCCA 271
Db 213 GTACGAGAGGAAAGATGGAACCTCAGTGTTCCTCAATGCAGACTAGATTACAGACGA 272
Qy 272 CAGNGGAGCCACGCGATCCGGGGGAGGAGCGACGACACTGATGCCGATGATGGTAG 331
Db 273 CAGGGGAGTCTCTGTTTGAAGGAGATGAAGATGAGGATGATGTTGATGATATCGAA 332
Qy 332 TGACTTCAACTACCTGCTGCTGGCACTGAGGACGAGAGATGCTGCACAGGAT 391
Db 333 TGAGTTCAATTACG-----CCAGGGAGCTTACACAGGC 365
Qy 392 GCGCAGCTGGCGATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAG 451
Db 366 GAGACACCAACGCCATGGCGAAGAGTTTCTTCTCTTAGACATGAATCTCAACCAAT 425
Qy 452 TGGCGAGATCGGCTCTCCAAGTATGACAGTGGAGAGATCCCTAGGGATACGTCCTTC 511
Db 426 TCCTTCTTCTCACCCTATGGCCATACGGTTTCTTGAGAGATTCGACGCTGTATACACAATC 485
Qy 512 AGTCACCAACAG-CCAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCATATGA 570
Db 486 TGTGCGAACTACATCAGGTCCTTTGGTCTCTTGCAGAGGATGCTATTTTCATCTCCATA 545
Qy 571 TGTCCCCCTACGGGGAACATCAGCAGACGCTGCTCCGTTTCCCTATGTGAATCATTCACCA 630

Db 546 TATTGATCCACGGCAACCTGTCCTGTAAAGAAATCGTGGACCCGTCAAAAGACTT----- 599
Qy 631 ATCCGTCAGGGAGTTCTCGCGCAGTATTGGGAATGTCCTCGGAAAGAGAGACTGTGATG 690
Db 600 -----GAACCTTATGGCGTTGGTAATGTTGACTGGAAAGAAAGAGTTGAAG 646
Qy 691 GCTGGAATAATGAAGCAGGACAAAGGGTGGGATTTCCCATGACTAATGGGCAAGCATTTGCTC 750
Db 647 CTTGGAAGCTGAAGCAGGAGAAAATATGTTACAGATGAC-----TGSTA 691
Qy 751 CCTCTGAAGTTCGGGAGCTACTGACATCGATGCACTACTCAATACACATGGAAGACG 810
Db 692 AATFACCATGAAGGAAGAGGAGAGAAAT---TGAAGGAGCTGTTCCCAATGGCGAAGAAC 748
Qy 811 CTTTACTGAATGAACACTGCGCAGCCCTCTATCTAGAAAAGTCCCATTTGCTTCCCTCA 870
Db 749 TCCAAATGGGTGATGATACACGCTTCTCCCTATGATGCTGCTGCTATCCCATCTTCTC 808
Qy 871 AAATAAATCCCTACAGAAATGCTATTGTTCTGCGGTTGCTTCTTAAGCATCTTCCCTGC 930
Db 809 GCCTAACCCCTTATCGGTTGCTGATTTATCTCGGCTTATCATCTGTGTTCTTCTTCTGC 868
Qy 931 ACTACGCTCAACAATTCCTGTCGCTAATGATGATACCATGCTGGGTTTATCTGTTATAT 990
Db 869 AATATCGTCAAACTCACCTGTGAAAATGCAATATCCTTTGTTGTTGACCTCGGTTATCT 928
Qy 991 GTGAGATTGGTTGCTTTATCTCGGATCTGATCAGTCCGAGTCCGAGTGGTTTCCCAATCA 1050
Db 929 GTGAGATCTGGTTGCTTTCTTGGCTTCTTGATCAGTTTCCCAATGGTACCCCATTA 988
Qy 1051 ACCGGGAGACCTACCTTGTATGACTGGCTTTAAGGTATGACCGAAGGTGAACGCTCTC 1110
Db 989 ACAGGGAGACTTATCTTGACGCTCTCGCTAATAGATATGATCGAGCGGTGAACCATCAC 1048
Qy 1111 AGTTGGCTGTGTTGACATATTTGTCAGTACAGTGCACCCCTTGAAGGAGCCACCTATCG 1170
Db 1049 AGCTGTTCTGTTGATGCTTTGTTAGTACAGTGCACCCCTTGAAGAGCCTCCCTTG 1108
Qy 1171 TCAGTGCACACTGTGCTATACATTTCTGTTGCTGTTGATTTATCCGTTGACAAAGTCTCTT 1230
Db 1109 TTACAGCAACACAGTTCTCTCGATTTCTTCTGTGGACTACCCGCTAGATAAAGTAGCCT 1168
Qy 1231 GCTATGATCTGATGAGGAGCTCAATGCTGACTTTTGACGCAATGCTGAGACTTTCAG 1290
Db 1169 GTTATGTTTACAGATGAGTTTACGCTATGCTTACCTTTGAAATCCCTTCTGNAACCGTG 1228
Qy 1291 AGTTTGTAGAAATGGGTACCATTTGTTGAAGATGATGACATTTGAACCCAGAGCTCCCG 1350
Db 1229 AGTTTGAAGAAATGGGTACCATTTTGAAGAAATTTCAACATTTGAACCTAGGGCCCTG 1288
Qy 1351 AGTTTGTACTTTGCCAGAAATTTGATTACTGAAAGACAAAGTCCAGGCTTCAATTTGTTA 1410
Db 1289 AATTCTATTTGCCAGAGATAGATTACTTGAAGCAAGATCCAAACCGCTTTTGTGA 1348
Qy 1411 AAGACCCCGGGCCATGAGAGAGATATGAAGATTTAAATACAGGATAAATGCCCTAG 1470
Db 1349 AAGAGCAGCAGGTATGAAGAGAGATATGAAGAGTTTAAAGTGAGGATAAATGCTCTTG 1408
Qy 1471 TTTCTAAGGCATTTGAAGTCCCCGAGGAAGGATGGATCATCAAGATGGCACACCATGGC 1530
Db 1409 TTGCCAAAGCACAGAAAATCCCTGAAGAGGCTGGACATGAGATGCTACTCCCTGGC 1468
Qy 1531 CAGAAACATATCAGGGATCATCTCCGGAATGATTCAGGTTTCTTGGTACAGTGGTG 1590
Db 1469 CTGGAACACACTAGAGATCATCTCGGAATGATACAGTGTTCTTAGGCCATAGTGGGG 1528
Qy 1591 GCCTTGTACTAGGGTATGAGCTCCCGGTTTATGTTTATGTTCTGCTGCTGAAAGCGTC 1650
Db 1529 GTCTGGATACCGATGGAATGAGCTGCCCTAGACTCATCTATGTTCTCTGTAAGAGCGC 1588
Qy 1651 CTGGTTCCAGCACCACAGAGGCTGGTCCCATGAAAGCCCTTGTGTTGCTCTCAGCTG 1710
Db 1589 CTGGATTTCAACACCACAAAGAGGCTGGAGCTATGAATGCAATTCGCTGATCTGCTG 1648

Qy 1711 TCCTTACTAATGCAACAATACAT 1770
Db 1649 TTCCTTACCNAATGGAGCATATCT 1708
Qy 1771 AGGCTGTCCGAGAACATATGTC 1830
Db 1709 AGGCTATTAAAGAAAGCTATGTC 1768
Qy 1831 ATGTGCAGTTCCCAACAAAGTT 1890
Db 1769 ATGTCCAGTTCCTCAACGTT 1828
Qy 1891 ACAGTGTCTTTTTTGTATATTA 1950
Db 1829 ATATAGTCTTTTTTCGATATTA 1888
Qy 1951 TGGGAACCTGTTGCTTTTCA 2003
Db 1889 TGGGTACTGTTGTTGTTTAA 1948
Qy 2004 ATTAAGCGGAAGAACGCCAGTT 2061
Db 1949 AAGAAGATTTAGAACCAAAATAT 2008
Qy 2062 GCAAGTCAAGAAAGAAAGGAGCTC 2121
Db 2009 AAAGTAGCAAGAAAGTAACTAG 2068
Qy 2122 TTCAGTATTCAATCTCGAAGAA 2181
Db 2069 CTCCACTTTTCAATATGAGGAG 2122
Qy 2182 AGAAATCAGTCTCTCATGCTCA 2241
Db 2123 AGAGGCTATTCTAATGTTCCAG 2182
Qy 2242 TTGTTGCCCTCCACTCTGATGGA 2301
Db 2183 TTATTGCGGCAACCTTTCATGGA 2342
Qy 2302 TTTTGAAGAAAGCTATCCATGTC 2361
Db 2243 TTCTGAAGGAGCTATTTCATGTC 2302
Qy 2362 CTGAGATTGGTTGGATCTATGGA 2421
Db 2303 AAGAGATTGGTTGGATCTATGTC 2362
Qy 2422 AGCAAGAGGCTGGCGTTTCAGTTC 2481
Db 2363 ATGCCCGGGTTGGATATCGATC 2422
Qy 2482 CCCCATCAATCTTTTCAGATCTG 2541
Db 2423 CACCAATCAATCTTCTTATGCTG 2482
Qy 2542 AAATCTTTTTCAGCGGCAATTCG 2601
Db 2483 AGATTCTCTTAGCAGACATGTC 2542
Qy 2602 TGGAGAGATTCCCTTACATCAAC 2661
Db 2543 TGGAGAGGATCGCTTATATCAAC 2602
Qy 2662 TCTATTGTATATGCTGCTATCT 2721
Db 2603 CGTATTGTATCTTCCCGCTTTT 2662
Qy 2722 GCAACTGCGCAGTATCTGGTCA 2781
Db 2663 GCAACTACCGAGTATTTGGTTCA 2722

Db 869 AATATCGTCAAACTCACCTCGTGAATAATGCATATCCTTGTGGTTGACCTCGGTTATCT 928
QY 991 GTGAGATTTGGTTGCTTTATCCTCGGATACGATCGCTCCGGAAGTGGTTTCCAATCA 1050
Db 929 GTGAGATCTGGTTTGCATTTCTTGGCTTCTTGATCAGCTTTCCCAATGGTACCCCAITA 988
QY 1051 ACCGGGAGACCTACCTTGTATGACATGGGTTTAAAGGTATGACCGAGAAAGGTGAACCGTCTC 1110
Db 989 ACAGGGAGACTTATCTTGACCTCTCGCTATGAAGATATGATCGAGACGGTGAACCATCAC 1048
QY 1111 AGTTGGCTGCTGTGACATATTTGTGACGTACAGTCGACCCCTTGAAGAGCCACTATCG 1170
Db 1049 AGTCGTTCTGTGTATGTTGTGTAGTACAGTGGGCCCAATGAAGAGCCTCCCTTG 1108
QY 1171 TCAGTCCCAACACTGTGCTATCCATTTCTGCTTGTGTATTCCTGCGTGACAAAGGTCTCTT 1230
Db 1109 TTACAGCAAAACACAGTTCTCTCGATTTCTGTGGACTACCCGCTAGATAAAGTAGCCT 1168
QY 1231 GCTATGATCTGATGACGAGCTTCAATGCTGACTTTTGACGCATTTGGCTGAGACTTCAG 1290
Db 1169 GTTATGTTTCAGATGATGGTTTCAGCTATGCTTACCTTTGAATCCCTTTCTGAAACCCGCTG 1228
QY 1291 AGTTTCTAGGAAATGGGTACCATTTGTGAAGAAGTATGACATTCGAACCCAGAGCTCCCG 1350
Db 1229 AGTTTCAAAAGAAATGGGTACCATTTTGAAGAAATTCACATTTGAACCTTAGGSCCCTG 1288
QY 1351 AGTTTACTTTTGGCAGAAATTTGATTTACCTTGAAGACAAAGTCCAGCTTCATTTGTTA 1410
Db 1289 AATCTTATTTTCCCGCAGAAATAGATTTACTTGAAGGACAAAGATCCAAACCGTCTTTTGT 1348
QY 1411 AAGACCGCGCGGCCATGAAGAGAAATATGAAGAAATTTAAATCAGGATAAATGCCCTAG 1470
Db 1349 AAGAGCGAGAGCTATGAAGAGAGATGAAGAGTTTAAAGTGAGGATAAATGCTCTG 1408
QY 1471 TTTCTAAGCATTGAAGTCCCGAGGAAAGATGGATCCGAAGTGGCAGACCATTCGC 1530
Db 1409 TTGCCAAAGCAGAAATTCCTGAAGAGGCTGGCAATGCAGGATGGTACTCCCTGGC 1468
QY 1531 CAGGAACAAATACCAGGATCATCTCGGAATGATTCAGGTTTCTTGGTCACAGTGGTG 1590
Db 1469 CTGGTAACAACACTAGAGATCATCTCGAATGATACAGGTGTTCTTAGGCCATAGTGGG 1528
QY 1591 GCCTTGATCTAGGATGATGACCTCCCGTTTGTAGTTGTGCTCTCGTGAAGAGCTC 1650
Db 1529 GTCTGATACCGTGAAGATGAGCTGCTAGACATCATGTTCTCGTGAAGAGGCG 1588
QY 1651 CTGGTTCCAGCACCAAGAGGCTGGTGCATGAATGCCCTTGTGCTGCTCAGCTG 1710
Db 1589 CTGGATTTCAACACCACAAAAGGCTGGAGCTATGAATGCATTCGCTGATCTGCTG 1648
QY 1711 TCCTTACTAATGCAATACATGTTGATCTTGATTTGATGATCACTACATCAACACAGCA 1770
Db 1649 TTCTTACCAATGGAGCATATCTTTTGAACGTGGATTTGTGATCTATTTAATAACAGTA 1708
QY 1771 AGGCTGTCGAGAAAGCTATGCTTCTTAATGGATCCAAAGCTAGGTCGCGCAAGTCTGTT 1830
Db 1709 AGGCTATTAAAGAAAGCTATGTTTCAATGATGGACCCGCTATTGGAAGAGTGTGCT 1768
QY 1831 ATGTGCAGTTCACCAAGAGTTTGAATGGATTTGATAGGAATGATCGATATGCAACAGGA 1890
Db 1769 ATGTCCAGTTCCCTCAACGTTTGTACGCTATTTGATTTGACGATTCGATATGCCAACAGGA 1828
QY 1891 ACACGTCTCTTTTGTATTAACCTTGAAGGCGCTTGAGCGCATTTCAAGGACCAAGTTATG 1950
Db 1829 ATATAGTCTTTTTCGATATTAACATGAAGGTTGGATGGTATCCAGGTCAGTATATG 1888
QY 1951 TGGGAACGTGTTGTTTTCACAGACAGCTATCTATGTTATGAGCC-----CCCA 2003
Db 1889 TGGGTACTGGTTGTTTGAATAGGCGAGGCTCTATATGGGTATGATGCTGTTTGTACGG 1948
QY 2004 ATTAAGGCGAAGAACCCAGGTTTCTTGGCATCA--CTATGTGGGGCGCAAGAAAGCAA 2061
Db 1948 ATTAAGGCGAAGAACCCAGGTTTCTTGGCATCA--CTATGTGGGGCGCAAGAAAGCAA 2061

Db 1949 AAGAAGATTTAGAACCAATAAT 2008
QY 2062 GCAAGTCAAGAAAGAGGAGCTC 2121
Db 2009 AAGTAGCAAGAAAGTATACAT 2068
QY 2122 TTCCAGTATTAATCTCGAAGAT 2181
Db 2069 CTCACCTTTTCAATATGAGGA 2122
QY 2182 AGAAATCAGTTCTCATGCTCA 2241
Db 2123 AAGGTCATTTCTAATGTCCTCA 2182
QY 2242 TTGTCCTCCACTCTGATGGA 2301
Db 2183 TTATTGCGCAACCTTTCATGGA 2242
QY 2302 TTTTGAAGAGCTATCATGCT 2361
Db 2243 TTCTGAAGAGCTATTCATGCT 2302
QY 2362 CTGAGATTTGGTTGGATCTAG 2421
Db 2303 AAGAGATTTGGTTGGATCTAG 2362
QY 2422 AGCAAGAGGCTGGCGTTTCAGT 2481
Db 2363 ATGCCCGGGTTGGATATCGAT 2422
QY 2482 CCCCCCAATCTTTTCAGATCG 2541
Db 2423 CACCAATCAATCTTTTCATCG 2482
QY 2542 AAATCTTTTCAGCGCGCATTCG 2601
Db 2483 AGATTTCTTTAGCAGACATTCG 2542
QY 2602 TGGAGAGATTCGCTTTACATCAAG 2661
Db 2543 TGGAGAGGATCGTTTATATCAAG 2602
QY 2662 TCTATTTGATATTCCTGCTATC 2721
Db 2603 CGATTTGATTTCTTCCCGCTTT 2662
QY 2722 GCAACTTGGCCAGTATCTGGTTC 2781
Db 2663 GCAACTACGCGAGTATTTGGTTC 2722
QY 2782 TTGAGATGAGTGGAGTGGTGTTC 2841
Db 2723 TGGAGCTGAGATGGAGCGGTGTC 2782
QY 2842 TCATTGGAGGTATCTCTGCACAT 2901
Db 2783 TCATTGGTGGACATCCGCCCAT 2842
QY 2902 CCGGTATCGACACCACTTCACAT 2961
Db 2843 CTGGTATCGACACCACTTCACAT 2902
QY 2962 CTGAGCTCTACATGTTCAAGTGG 3021
Db 2903 CAGAACTCTACATCTTCAATGG 2962
QY 3022 TTAACATGTTGGTGTCTGTTGCT 3081
Db 2963 TGAACCTCATAGGCAATTTGGCT 3022
QY 3082 GGGGCCGCTCTTTGGGAAGCTCT 3141
Db 3023 GGGGTCCGCTTTTCGGGAAGCTCT 3082

QY 1291 AGTTTGTAGAAATGGGTACCATTTGTGAGAAAGTATGACATTTGACCCAGAGCTCCCG 1350
DB 1299 AGTTTGCACAAAGAAATGGGTACCATTTTGCAGAAATTTCAACATTTGAACCTTAGGGCCCGCTG 1358
QY 1351 AGTTTACTTTTGGCCAGAAATTTGATTACCTTGAAGACAAAGTCCAGCCCTTCATTTGTTA 1410
DB 1359 AATTCATTTTGGCCAGAAATGATATTTGTTGAAGACAAAGATCCACCGCTCTTTGTTA 1418
QY 1411 AAGACCGCGGGCCATGAAGAGAGAAATATGAAGAAATTTAAATCAGAGATAAATGCCCTAG 1470
DB 1419 AAGACGACAGCTATGAAGAGAGAGATGAAGAGATTTAAAGTGAAGATAAATGCTCTTG 1478
QY 1471 TTTCTAAGGCATTTGAAGTCCCGAGGAAGATGATCATGCAAGATGGGACACACCATGGC 1530
DB 1479 TTGCCAAGCACAGAAATCCCTGAAGAGGCTGGCAATGCAAGATGGTACTCCCTGGC 1538
QY 1531 CAGGAACAATPACAGGAGTATCCTGGAATGATTCAGGTTTTCCTTTGTPCACAGTGGT 1590
DB 1539 CTGGTAAACACACTAGAGATCATCTGGAATGATACAGGTGTTCTTAGGCCATAGTGGG 1598
QY 1591 GCCTTGATACAGGTAAGTAAATGAGTCCCGCTTTCCTGTTATGTTGCTCGTGAAGCGTC 1650
DB 1599 GTCTGGATACGATGGAATGAGTGCCTAGACTCATCTATGTTTCTCGTGAAGCGGC 1658
QY 1651 CTGGGTTCCAGCACCAAGAGGCTGGTCCATGAATGCCCTTGTTCGTCTCAGCTG 1710
DB 1659 CTGGATTTCAACACACCAAAAGGCTGGAGTATGAATGATCCCGTGTATCTGTTG 1718
QY 1711 TCCTTACTAATGAGCAATACATGTTGAATCTGATTTGATGATCACTACATCAACAACAGCA 1770
DB 1719 TTCCTACCAATGGAAGATATCTTTTGAACGTGGATTTGATCAATTTACTTTAATAACAGTA 1778
QY 1771 AGGCTGCCGAGAGCTATGCTTCTTAATGGATCCAAACCTAGGTCGCAAGTCTGTT 1830
DB 1779 AGGCTAATAAAGAGCTATGTTTTCATGATGGACCGGCTATTGGAAGAAGTGTCTGCT 1838
QY 1831 ATGTGCAAGTCCCAAGAGTTGATGGATTTGATAGGATGATAGATGATGATGCAACAGCA 1890
DB 1839 ATGTCCAGTCCCTCAAGCTTTTACCGTATGATTTGACGATGATATGCAACAGCA 1898
QY 1891 ACATGCTTTTTTATTAACCTTGAAGGCGCTTACGCGCATCAAGACCAAGTATTATG 1950
DB 1899 ATATAGTCTTTTCGATTAATCAATGAAGGGTTGGATGATCCAGGGTCCAGTATATG 1958
QY 1951 TGGCACTGTTGTTTTCACAGAACACCTATCTATGTTATGAGCC-----CCCA 2003
DB 1959 TGGTACTGTTGTTTATTAAGCAGGCTCTATATGGGTATGATCCTGTTTGGACGG 2018
QY 2004 ATTAAGCGCAAGAGCCAGGTTTCTTGGCATCA--CTATGTGGGGCAAGAAGGCA 2061
DB 2019 AAGAAGATTTAGAACCAATATATTGTCAGAGCTGTTGCGGGTCAAGGAAGAAGGTA 2078
QY 2062 GCAAGTCAAGAAAGAGCTCAGATAAGAAAAGTGAACAAAGCATGTGGACAGTTCTG 2121
DB 2079 AAAGTAGCAAGAAATTAACATAGCAAAAGAGAGAGCATCAACAAGAGTGACTCCAATG 2138
QY 2122 TTCCAGTATTAATCTCGAAGCATAGAGGAGGTGTTGAAGTGTGGTGTGATGATG 2181
DB 2139 CTCCACTTTCAATATGGAGCATCATGAGGGTTTGAAGGT-----TATGATGATG 2192
QY 2182 AGAAATCAGTTCTCATCTCAATAGCTTAGAGAGATTTGGCCAGTCAGCAGCAT 2241
DB 2193 AGAGGTCTATTCTAATGTCACAGAGAGTGTAGAGAAGCGTTTGTGTCAGTCGCGGTAT 2252
QY 2242 TTGTTGCTCCACTCTCATGATATGGTGTTCCTAGTCTCCACTCCAGATCTC 2301
DB 2253 TTATTGGGCAACCTTCATGGAACAGGCGCATTCACCAACAACCAATCCGCTACTC 2312
QY 2302 TTTTGAAGAAGCTATCCATCTCATAGTTGCTGCTTAGGACAAAGTCTGAATGGGAA 2361
DB 2313 TTCTGAAGAGGCTATTCTATGTTTATAGCTGTGTTACGAGACAAAGATGATGGGCA 2372
QY 2362 CTGAGATGGTTGATCTATGATCTGCTCAGAGAAGATATCTTACTGATGATCAAGATGC 2421

DB 2373 AAGAGATTGGTTGGATCTATGG 2432
QY 2422 AGCAAGAGGCTGGCGTTTCAGT 2481
DB 2433 ATGCCCGGGTTGGATATCGAT 2492
QY 2482 CCCCCCAATCTTTTCAGATCG 2541
DB 2493 CACCAATCAATCTTTCTGATCG 2552
QY 2542 AAATTTCTTTTCAGCGCGCATTCG 2601
DB 2553 AGATTTCTTTTAGCAGACATTCG 2612
QY 2602 TGGAGAGATTCGCTTACATCA 2661
DB 2613 TGGAGAGATCGCTTATATCA 2672
QY 2662 TCTATTGTATATTCCTGCTAT 2721
DB 2673 CGTATGTATTTCTCCCGCTTT 2732
QY 2722 GCAACTTGGCCAGTATCTGGTT 2781
DB 2733 GCAACTACGCGAGTATTGGTT 2792
QY 2782 TTGAGATGAGTGGAGTGGTGT 2841
DB 2793 TGGAGCTAGATGGAGCGGTGT 2852
QY 2842 TCATTGGAGGTATCTCTGCACA 2901
DB 2853 TCATTGGTGACATCCGCCCA 2912
QY 2902 CCGGTATCGACACCACTTCAC 2961
DB 2913 CTGGTATCGACACCACTTCAC 2972
QY 2962 CTGAGCTTACATGTTCAAGTG 3021
DB 2973 CAGAATCTACATCTCAATG 3032
QY 3022 TTAACATGGTTGGTGTCTGCT 3081
DB 3033 TGAACCTCATAGGCAATTTGGC 3092
QY 3082 GGGGCGCTCTTTGGGAAGCT 3141
DB 3093 GGGTCCGCTTTTGGGAAGCT 3152
QY 3142 TCCCAAGGCTTTATGGGCA 3201
DB 3153 TCTTGAAGGTCTCTTGGGA 3212
QY 3202 TCCCTCTGCTCTATCTCTC 3250
DB 3213 TCTTCTGCGCTCCATCTTCTC 3261

RESULT 14

AX030962

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplan

Spermatophyta; Magnol

Rosidae; eurosids II

REFERENCE

1 (bases 1 to 3673)

573 bp DNA linear PLN 20-SEP-2000
09000549.Streptophyta; Embryophyta; Tracheophyta;
yta; eudicotyledons; core eudicots;
Brassicaceae; Arabidopsi

QY 1651 CTGGTTCCAGACCAACAGAGCGTGGTCCGATGAATGCCCTTGTTCGTGCTCAGCTG 1710
DB 1659 CTGGATTCAACACACACAAAGCGTGGAGCTATGATGCATTTGATCGTGTATCTGTTG 1718
QY 1711 TCCTTACTATGACAAATACATGTTGAATCTTGATTTGATCAGTACATACAAACAGCA 1770
DB 1719 TTCTTACCAATGAGGATATCTTTTGACGTGATTTGATCATTTACTTTATACAGTA 1778
QY 1771 AGGCTGTCCGAGAGCTATGTTGCTTCTTAATAGGATCCAAACCTAGGTCCGCAAGTCTGTT 1830
DB 1779 AGGCTATRAAGAAGCTATGTTTTCATGATGACCGCGCTATTGGAAGAAGTCTGCT 1838
QY 1831 ATGTGAGTTCCACAAAGTTTGATGGATTTGATAGGATGATCGATATGCAACAGGA 1890
DB 1839 ATGTCCAGTTCCCTCAACGTTTGGCGTATTTGATTTGACATCGATATGCCAACAGGA 1898
QY 1891 ACAGTCTGTTTTTGTATTAATTAAGTGTGAGGGCTTTGAGGGCATTCAGAGGACCAAGTTATG 1950
DB 1899 ATATAGTCTTTTCGATATTAACATGAAGGGTGTGATGGTATCCAGGGTCCAGTATATG 1958
QY 1951 TGGGAAGTGTGTTGTTTCAACAGACAGCTATCTATGTTGTTATGAGCC-----CCCA 2003
DB 1959 TGGGTACTGTTGTTTAAATAGGACAGGCTATATGTTGGTATGATCCTGTTTGTAGCG 2018
QY 2004 ATTAAGCGGAAGACGAGTTTCTTGGCATCA--CTATGTGGGGGCAAGAAGAGGCAA 2061
DB 2019 AAGAAGATTTAGAACCAATATATTGTCAGAGCTGTTGCGGGTCAAGGAAGAAAGTA 2078
QY 2062 GCAAGTCAAGAAAGAGGCTCAGATAGAAAAAGTCAACAGAGCTGTGACAGTCTG 2121
DB 2079 AAGTAGCAAGAAGTATTAAGTACGAAAGAGGAGGAGGATCAACAGAAGTACCAAT 2138
QY 2122 TTCCAGTATTCAATCTCGAAGACATAGAGAGGGTGTGAAGGTGCTGGGTTTGATGATG 2181
DB 2139 CTCACATTTTCAATATGGAGGACATCGATGAGGGTTTGAAGT-----TATGATGATG 2192
QY 2182 AGAATCAGTTCATGCTCAANTGAGCTTAGAAGAGATTTGGCCAGTCAGCAGCAT 2241
DB 2193 AGAGGTCTATTCTAATGTCCTCAGAGAGGTGTAGAAAGCGTTTGGTCAGTCCGCGGTAT 2252
QY 2242 TTGTTGCCCTCCACTGTATGGAATATGTTGTTTCCCTCAGTCTCTCCACTCCAGAAATCTC 2301
DB 2253 TTATTCGGGCAACCTTCATGGAAGAGGCGGCATTCACCAACACCAATCCCGCTACTC 2312
QY 2302 TTTTGAAGAAGCTATCCATGTCATAGTTGTGGCTATGAGGACAAAGTCTCAATGGGAA 2361
DB 2313 TTCTGAAGGAGGCTATTCATGTTAAGCTGTGGTTACGAAGCAAGACTGAATGGGSCA 2372
QY 2362 CTGAGATTGGTTGGATCTATGGATCTGTACAGAAAGATATCTTACTGGATTCAAGATGC 2421
DB 2373 AAGAGATTGGTTGGATCTATGTTCCGTGACGGAAGATATCTTACTGGTTCAAGATGC 2432
QY 2422 ACCGAGAGGCTGGCGTTTCAGTCTATTCGATGCCCAAGCGCCAGCTTTCAAGGATCTG 2481
DB 2433 ATCCCGGGGTTGGATATCGATCTACTGCAATCCCTCCACGCCCTCGCTTCAAGGATCTG 2492
QY 2482 CCCCCATCAATCTTTTCAGATCTGTGAACCAAGTCTCGGGTGGGCTCTCGGTTCTGTTG 2541
DB 2493 CACCAATCAATCTTTCTGATCTTTGAACCAAGTCTTTCGATGGCTTTGGATCTATCG 2552
QY 2542 AATTTCTTTTCAGCGGATTTGCCCTTATGTTATGGTATGCTAGGAGGGCGCTCAAGTTC 2601
DB 2553 AGATTCTTTTACGACAGATATCTCTATCTGTTATGTTTACCATGGAAAGTTGAGACTTT 2612
QY 2602 TGGAGAGATTCGTTTACATCAACACCAAGCTATTACCCACTACCTCTCTCCGCTTCTAG 2661
DB 2613 TGGAGAGGATCGTTTATATCAACACCAAGCTATCTATCTTATACATCATCCCTCTTATG 2672
QY 2662 TCTATTGTATGCTGCTATCTCTGCTCAGTGGAAAGTTTCTATCATGATGCCAGATTA 2721
DB 2673 CGTATTGTATTCTCCCGCTTTTGTCTCTCATCACCAGCAGATTTCTATATACCCGAGATA 2732

QY 2722 GCAACTTGGCAGTATCTGGTT 2781
DB 2733 GCAACTACCGAGTATTGGTT 2792
QY 2782 TTGAGATGAGGTGAGGTGTT 2841
DB 2793 TGGAGCTGAGATGGAGCGTGT 2852
QY 2842 TCATTGGAGGTATCTCGCACA 2901
DB 2853 TCATTGGTGGACATCCGCCCA 2912
QY 2902 CCGGTATCGACACCAACTTCAC 2961
DB 2913 CTGGTATCGACACCAACTTCAC 2972
QY 2962 CTGAGCTCTACATGTTCAAGTG 3021
DB 2973 CAGAACTCTACATCTCAATG 3032
QY 3022 TTAACATGTTGGTGTCTGTC 3091
DB 3033 TGAACCTCATAGGCATTTGGCT 3092
QY 3082 GGGGGCGCTCTTTGGGAAGCTG 3141
DB 3093 GGGTCCGCTTTTCGGGAAGCTG 3152
QY 3142 TCCTCAAGGTTCTTATGGCAGG 3201
DB 3153 TCTTGAAGGTTCTGTTGGGAAG 3212
QY 3202 TCTCTCTGCTCTCTATCTTCTCC 3250
DB 3213 TTCTTCTCGCTCTCATCTCTCG 3261

RESULT 15
AF200532
LOCUS
DEFINITION
Zea mays cellulose synthase-8 (Cesa-8) mRNA, complete cds.
ACCESSION
AF200532
VERSION
AF200532.1 GI:962288
KEYWORDS
Zea mays.
SOURCE
Zea mays.
ORGANISM
Eukaryota; Viridiplaneta; Spermatophyta; Magnoliopsida; clade: Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 3812)
Holland, N., Holland, D., Xoonostle-Cazares, B., Delmer, D.P., Dhugga, K.S., and Dentjaris, T., 1999, The plant cellulose synthase (Cesa) gene family
JOURNAL
Plant Physiol. 123 (4): 13-1324 (2000)
MEDLINE
20398328
PUBMED
10938350
REFERENCE
2 (bases 1 to 3812)
Dhugga, K.S. and Helen, S., T.G.
AUTHORS
Direct Submission
TITLE
Submitted (01-NOV-1999)
JOURNAL
International, Inc., 2000
FEATURES
Location/Qualifiers
1. .3812
/organism="Zea mays"
/db_xref="taxid:577"
/note="identified and isolated from the maize genome collection of Pioneer Hi-Bred International, Inc."
1. .3812
/gene="Cesa-8"
214. .3498
/gene="Cesa-8"

/note="CESA-8"
/codon_start=1
/product="cellulose synthase-8"
/protein_id="AAF89968.1"
/db_xref="gi:9622888"
/translation="MEASAGLVAGSHNRNELLVIRRDRESGAAGGAARRABAPQIC
GDEVGVGDGPFVACNECAFPVACACIYERREGSOAQPOCTRYKRLKGPVRVAGD
BREDVDLEGEFLQDAAHEDDPOVVAESMLRAOMSYGRGDHAGPSPVNPVPL
TNGMVDIDPEQHALVPSYMSGGGGKKRIHPLPADNPLVOPSPMDPSKDLAAYG
YGSVAWKEGMEGWKQERLQHVRSSEGGDWDADLPLMEARQPLSRKVPISSSR
INPYRMILIVRLVIGEFHYRVMHPAKDAFALWLLSVICEIWFAMSWILDQFPKWL
IERETYLRLDFDKGQSLAPIDFVSTVDPTEKPLVNTANTVLSILSDVPE
KYSVSDGAAMLTFEALSETSEFAKKWPFKEINIPRAPWTFQOKIDYLDKV
AASFVREMRAREVEBEFKVINALVAKQVPEEGTWQDGPWPNVNRDHPQMIQ
VFLGSGDREGNELPLRVVYSREKRPYNNHKKAGAMNALVRYSAVLSNAYLLNL
DCDHYINNSKAIKEMCFMDPLGVKVCYVQFQBEDGLDKNRDANRVNVEFDNM
KGLDGIQPIVVGTCVFRQALYGDAPTKKPPSPRTNCWPKWLSLSCCSNRNKK
KITPKTEKKRRLFFKKAENPSPAYALGEIDEGAPADIEKAGIVNQKLEKFGOSS
VFVASTLLENGTLKSPASLSLEKAIHVISCGYEDKTDWKEIGWIYGSITDILTG
FKMHGWRSTYICPKRPAFGSAPLNSDLRLHQVLRWALGSVEIPFSKPCWLYGY
GGLKPLRFERSINSIVPWTSPILLAYCTLPALCLLTGKFIPELTNVASIFWMAFI
CISVTGILDEMRWSGVAIDWNRNQFWIGVSAHLFAVFOGLLKVFAGIDTSFTVTS
KAGDDEESELVTEKMTLLIPTPTLLLFNFIQVAGISNAINNGYSWGLFGKLF
AFWVTHVLPFLKGLVGRQRTPIVLIWSILLASIFSLNVRDPLAKSNPLLEE
CGLDCN"

BASE COUNT 873 a 966 c 1111 g 862 t

ORIGIN

Query Match		37.0%; Score 1340.2; DB 8; Length 3812;
Best Local Similarity		65.5%; Pred. No. 1e-253;
Matches 2137; Conservative		0; Mismatches 1023; Indels 103; Gaps 8;
QY	89	CGGAGGACGCGGGCGGAGCGTGTGCAGATCTGCCCGCAGCGCGTGGCACCAGGTT 148
DB	309	CGCGGGCGCGGGCGGAGCGCGTGTGCAGATCTGCCCGCAGCGCGTGGGCGTT 368
QY	149	GGAGCGGACGCTTTCACCGCGTGGAGCGTCTCCCGCTTCCCGCTGCCCGCCCTGCTA 208
DB	369	CGACGGGAGCGCTTCTGTCGCGTGCACAGAGTGCCTTCCCGCTGCCCGCGCTGCTA 428
QY	209	CGAGCAGGCGGAGGAGGACCGAGCGTGTGCAGATCTGCCCGCAGCGCGTGGCACCAGCG 268
DB	429	CGAGTACGAGCGCGGAGGCGTGCACAGCGTGTGCCCGCAGTGCAGGACCGCGCTACAGCG 488
QY	269	CCACAGAGGAGCCACGCGATCCGGGGGAGGAGGCGGACGACACTGATGCCCGCTATGG 328
DB	489	CCTCAAGGCGTCCCGCGGTGGCGGACGAGGAGGAGCGCGTGCAGGACCTGGA 548
QY	329	TAGTGACTTCAAC---TACCTGATCTGGCACTGCGACTGAGACAGAGAGAGATTGCTGA 385
DB	549	GGCGGAGTTGGCGCTCAGGACGCGCGCGCCACAGGACGACCGCGAGTACGTCGCCGA 608
QY	386	CAGGATCGCAGCTGGCGCATGAACACCGGGGGCAGTGGCAATGTGGCCACCCCAAGTA 445
DB	609	GTCCATGCTCAGGCGCCAGATGAGCTACGGCGCGCGCGCGCGACCGCGCTTCAG 668
QY	446	TGACAGTGGGAGATCGGCTCTCCAAATATGACAGTGGAGAGATCCCTAGGGGATACGT 505
DB	669	CCCGCTCCCAACGCTCCCTCCTCACCACACG-----GCCAGATGGTTGAT 714
QY	506	CCCTTCAGTCAACCAACGAGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCA 565
DB	715	GACATCCCGCGGAGCAGCAGCGCTGCTGCGCTCTACATGAGCGGCGGCGGGCGG 774
QY	566	CATGATGTCCCTTACGGGGAACATCAGCAGAGCGTGTCCGTTTCCCTATGTGAATATTC 625
DB	775	GGCAAGAGGATCCACCCGCTCCCTTTCCGAGATCCCAACCTCCAGTGCACACGAGATCC 834
QY	626	ACCAAAATCCGCTCAAG--GAGTCTCCGGCAGTATTTGGAATGTGCTCGAAGAGAGA 683
DB	835	ATGACCCGCTCAAGGATCTGCGCGCTACGGATATGCGACGCTGGCTTGGAGGAGAGA 894
QY	684	GTTGATGGCTGGGAAAATGAAGCAGGACAAAGGCTGCGATTCCTCATGACTTAATGGACAAGC 743

DB	895	ATGAGGGCTGGGAAGCAGAGCAGGAGCGCTGCAGCATGTGCAGGAGCGAGGG----- 947
QY	744	ATTGCTCCCTCTGAAGGCTGGGCGAGCTACTGACATCGATCGATCTACTGATPACAACTAG 803
DB	948	-----TGGCGGTGATTTGGGATGGCAGCAT 972
QY	804	GAAGACGCTTTACTGAATGATGAAACTCGCCAGCCCTCTATCTAGAAAAGTCCCCATCTCT 863
DB	973	GCAATCTGCCCATTAATGGATGAAGCTAGGACGCCATTTCCAGAAAAGTCCCTATATCA 1032
QY	864	TCCTCCAAAATAAATCCCTACAGAAATGGTATTTCTCGCGTGTGGTTTCTTAAGCATC 923
DB	1033	TCAAGCGCAATTAATCCCTACAGGATGATATCGTTATCCGTTATCCGTTGTGGTTTTC 1092
QY	924	TTCTGCACTACCTCTCACAAATCCTGTGCGTAAATGATACCCACTGTGGCTTTTATCT 983
DB	1093	TTCTTCCACTACCGAGTGATGCAATCCGGGGAAGATTTGCAATTTGGCTTGCCTATATCT 1152
QY	984	GTTATATCTGAGATTTTGGTTTGGCTTTTCTTCTCGATCTGATCAGTTCCTCCGAAGTGGTT 1043
DB	1153	GTAATCTGTGAATCTCTGTTTGGATGTCCTGGATTTCTTGTATCAGTTCCCAAGTGGCTT 1212
QY	1044	CCAATCAACCGGAGAGACTTACCTTGATAGACTGGCTTTTAAAGTATGACCGAGAAAGTCAA 1103
DB	1213	CCAATCGAGAGAGAGACTTACCTGGACCGTTTGTCACTAAGGTTTGCAAGGAGGTCAA 1272
QY	1104	CCGCTCAGTGTGGCTGTGTGACATATTTGTCACTACAGTGCAGCCCTTTGAAGGAGCA 1163
DB	1273	CCCTCAGCTTGTCTCCAATCGACTTCTTGTCACTACGTTTGTATCCCAAGAACTCT 1332
QY	1164	CCTATCTGACTGCCACACTGTGCTATCCATCTTGTCTGCTTGTGATTTCCCGTGACAA 1223
DB	1333	CCCTTGTGTCAGGAACTGCTCTTTCATCTTCTTGTGGATTTATCCGTTTGGAGAG 1392
QY	1224	GTCTCTTGTATCTGATGACGAGGCTTCAATCTGACTCTGACTTTTGAAGCATTTGCTGAG 1283
DB	1393	GTCTCTGCTATGTTCTGATGATGCTGCTGCAATCTTACGTTTGAAGCATTTGCTCAA 1452
QY	1284	ACTTCAGAGTTGCTAGGAATGGTACCAATTTGTGAAGAAGTATGACATGAACCCAGA 1343
DB	1453	ACATCTGAATTTGCAAGAATGGTTCCTTCCAGCAAAAGTTTAATATCAGGCTCGT 1512
QY	1344	GCTCCGAGTTTACTTTTCCAGAAAATTTGATTTACTGAAAGACAAAATCCAGGCTTCA 1403
DB	1513	GCTCTGAGTGGTACTTCCAAACAGAGATAGACTACCTGAAAGACAAGGTTCTCTCTCA 1572
QY	1404	TTTGTAAAGACCCCGGCCATGAAGAGAGATATGAAGAATTTAAATCAGGATTAAT 1463
DB	1573	TTTGTAGGGAGAGGCGGCGATGAAGAGAGAAATCAGAGAAATTCAGGATGAAGTCAAT 1632
QY	1464	GCCCTAGTTTCTAAGGCATTTGAAAGTCCCGGAGAAAGTGGATCATCAAGATGGCACA 1523
DB	1633	GCTTGTGTGCAAAAGCCCAAAAGTTTCTTGAGAAAGTGGACAATGCAAGATGGAAGC 1692
QY	1524	CCATGGCCGCAACATACCGGGATCATCTCGGAATGATTTCAAGTTTCTTGGTCAAC 1583
DB	1693	CCCTGGGCTGGAACACGCTACGCGATCATCTTGGAAATGATTCAGGATTTCTTGGCCAA 1752
QY	1584	AGTGGTGGCTTGATCTAGGGTAAATGAGCTCCCGCTTTAGTTTATGTGCTCGTCAA 1643
DB	1753	AGTGGGCTGCTGATGTGGAAGAAATGAGTTGCTCGCTGCTTATCTCTCGAGAGAA 1812
QY	1644	AAGGCTCTGGTTCAGCACCACAGAGGCTGGTGGCATGATGATGCCCTTGTCTGCTGC 1703
DB	1813	AAGAGCCAGGTTTAAACCATCACAGAGGCTGGTGGCATGAATGCATGGTCTGGTCTC 1872
QY	1704	TCAGTCTCTTACTTAATGACCAATACATGTTGAATCTTGTATTTGATGACTCACATCAAC 1763
DB	1873	TCTGCTCTTATCAATGCTGCATACCTATTGAATCTGACTGTGATGATCACATCAAC 1932
QY	1764	AACAGCAGGCTGTCCGAGAGGATGTGCTTCTTAATGGATCCAAACCTAGGTCGCGAA 1823

Db 1933 AATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTGGGAAGAAA 1992
Qy 1824 GTCTGTTATGTGAGTCCACAAAAGTTTGTATGGATGATAGGAATGATGCTGATATGCA 1883
Db 1993 GTGTGCTATGTACAGTTCCTCAGAGTTTGTATGTATGACAAAATGATCGATACGCT 2052
Qy 1884 AACAGGAACACTGTCTTTTGTATATAACTTGTAGGGCCCTTGACGGCATCAAGGACCA 1943
Db 2053 AACAGGAACCTTGTCTTTTGTATATAACTTGTAGGGCCCTTGACGGCATCAAGGACCA 2112
Qy 1944 GTTTATGTGGAACTGTGTGTTTCAACAGAACAGCTATCTATGGTTATGAGCCCCCA 2003
Db 2113 ATTTATGTGGTACTGTATGTTTTCAGACGGCAGCTATGTTTATGATGCTCTCT 2172
Qy 2004 ATTAAGCGGAAGAACCA-----GGTTTCTTGGCATCA 2036
Db 2173 AAAACGAAGAAGCCACATCAAGAACTTGCAACTGTGSCCCCAAGTGGTCTCTCTTGC 2232
Qy 2037 CTATGTGGGGCAAGGAAGCAAGTCAAGTCAAAAGAAAGAGCTCAGATAAGAAAAA- 2095
Db 2233 TGCTGCAGCAGGAACAAGATAAAAGAAAGACTACAAAACAAAGACGGAGAAAGAAA 2292
Qy 2096 -----GTCCGAACAGCATGTGACAGTCTGTGTCCAGTATTCATCTCGAAGACATAGAG 2150
Db 2293 AGATTATTTTCAAGAAAGCAGAAAACCCATCTCTGCTATATGCTTTGGGTGAATTTGAT 2352
Qy 2151 GAGGGTCTTGAAGTGTCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGC 2210
Db 2353 GAAGGTCTCCAGGTGC-----TGATATCGAAGAGCGCGGAATCGTAAATCAACAGAAA 2406
Qy 2211 TTAGAAAGAGATTGGCCAGTCAGCAGCATTTGTGTGCTCCACCTCTGATGGAATATGTT 2270
Db 2407 CTAGAGAAGAAATTTGGGAGTCTTCTGTTTGTGCGATCAACACTTCTTGAGAACGGA 2456
Qy 2271 GGTGTTCTCAGTCCCTCCACTCCAGAACTCTTTTGAAGAGACTATCCATGCTCATAGT 2330
Db 2467 GGNCCCTGAGAGCCAGTCCAGTCTCTCTTGAAGAGACTATACATGTTATACAG 2526
Qy 2331 TGTGGCTATGAGGACAAAGTCTGAATGGGAAGTGTGATGTTGGATCTATGGATCTGTC 2390
Db 2527 TGGGGCTACGAAGCAAGACCCAGTGGGGAAGAGATGGGTGGATTTACGGATCGATC 2586
Qy 2391 ACAGAGATATCTTACTGATTCAGATGCACGCAAGAGGCTGGCTCAGTCTATATGC 2450
Db 2587 ACAGAGATATCTTACTGATTCAGATGCACGCAAGAGGCTGGCTCAGTCTATATGC 2646
Qy 2451 ATGCCAAGGCCAGCTTCAAGGATCTGCCCCATCAATCTTTCAGATCGTCTGAAC 2510
Db 2647 ATCCGAAGGCCAGCTTCAAGGATCTGCCCCATCAATCTTTCAGATCGTCTGAAC 2706
Qy 2511 CAAGTGTGGGGTGGCTCTCGGTTCTGTGAAATCTTTTCAGCGCGCATGGCCCTTA 2570
Db 2707 CAGGTCTTGGTGGCCCTTGGGTCGTCGAAATTTCTTCAGCAAGCACTGCCACTT 2766
Qy 2571 TGTATGGCTACGGAGGGCCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACC 2630
Db 2767 TGTATGGGATACGGCGGGGCTAAAAATCTCTGGAAGGTTTCTTATATCAACTCCATC 2826
Qy 2631 ATTTACCACTAACCTCTCTCCGCTCTAGTCTATGATATGCTCCCTGCTATCTGCTG 2690
Db 2827 GTTTATCCCTGGACGTCCATCTCTCCGCTTACTGTACCTTGCCTGCCATCTGCCCTG 2886
Qy 2691 CTCACGTGAAGTTCATGTCAGAGATAGCAAGTTCAGAGATTCGCTTACATCAACACCACC 2750
Db 2887 CTCACGGGAAGTTTATACACACAGAGCTTACCAATGTCCGAGTATCTGTTTCAATGCG 2946
Qy 2751 CTCCTCTTCAATTTCCCACTGCTATCTTGAAGAGTGGTGGAGTGGTGGCATT 2810
Db 2947 CTTTTCACTGCACTCCCGGACCCGATCTCTGGAATGAGGTGGAGTGGCGGCCATC 3006
Qy 2811 GACAGTGGTGGAGGAATGAACAGTCTGGGTCAATGGAGGTATCTCTGCACATCTGTTT 2870
Db 3007 GACGACTGGTGGAGGAACGAGAGTCTGGGTCTATCGGAGGGTTCGGCGCATCTGTTTC 3066

Qy 2871 GCGTCTTTCCAGGGTCTTCTGA 2930
Db 3067 GCGGTCTTCCAGGGCTGTGTA 3126
Qy 2931 TCAAGGGCTAATGACGAAGAAG 2990
Db 3127 TCAAGGGCTGTAACGTTTCAAGTGGACCA 3183
Qy 2991 CTCTCATCCCTCCGACGACCA 3050
Db 3184 CTGCTGATACCCCGACCAAGC 3243
Qy 3051 TCCTACGCCATCAACAGTGGT 3110
Db 3244 TCAAGCGCATCAACAAACGGT 3303
Qy 3111 GCCTTCTGGGTGATTTTTCAT 3170
Db 3304 GCCTTCTGGGTGATCTCCACC 3363
Qy 3171 CGCACACCGACGATTTGTCATC 3230
Db 3364 AGGACCGCGACGATCTGTCATC 3423
Qy 3231 TGGGTTCTGTTGATCCATTTCA 3290
Db 3424 TGGGTCGGCTGACCCGTTCC 3483
Qy 3291 ATCAACTGTAGGAAGTGGGA 33
Db 3484 CTGGACTGCAACTGAAGTGGG 36

Search completed: February 25, 2003 18:20:39
Job time : 9088 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 14:53:29 ; Search time 694 Seconds
(without alignments)
11766.204 Million cell updates/sec

Title: US-09-900-237-29

Perfect score: 3626

Sequence: 1 gcacgaggaaccccgctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2595	71.6	3799	24	AA16458
2	2572.2	70.9	3746	21	AA299512
3	2572.2	70.9	3746	21	AA299527
4	2572.2	70.9	3773	21	AA299494
5	2325.2	64.1	3704	21	AA299533
6	1681.6	46.4	1734	21	AA258271
7	1677.6	46.3	3198	21	AAC49550
8	1673.6	46.2	3614	19	AAV06567
9	1408.4	38.8	3725	21	AA299500

10	1408.4	38.8	3725	21	AA299506	DNA encoding a mai
11	1408.4	38.8	3725	21	AA299521	DNA encoding a mai
12	1382	38.1	3851	21	AA671114	Pinus radiata cell
13	1373	37.9	3753	21	AA299515	DNA encoding a mai
14	1373	37.9	3753	21	AA299530	DNA encoding a mai
15	1373	37.9	3780	21	AA299497	DNA encoding a mai
16	1343.4	37.0	3603	19	AAV06565	Arabidopsis cellu
17	1340.2	37.0	3786	21	AA258265	Corn cellulose syn
18	1340.2	37.0	3813	21	AA299509	DNA encoding a mai
19	1340.2	37.0	3813	21	AA299524	DNA encoding a mai
20	1338.6	36.9	3673	19	AAV06568	Arabidopsis cellu
21	1331.4	36.7	2830	24	AA516455	Corn cDNA encoding
22	1320.4	36.4	3776	21	AA258263	Corn cellulose syn
23	1317.2	36.3	3568	21	AA299491	DNA encoding a mai
24	1271	35.1	3936	21	AA258266	Corn cellulose syn
25	1271	35.1	3969	21	AA299503	DNA encoding a mai
26	1271	35.1	3969	21	AA299518	Soybean cellulose
27	1257	34.7	3517	21	AA258268	Cotton cellulose s
28	1217.2	33.6	3328	19	AAV34432	Cellulose synthase
29	1212.4	33.4	3207	19	AAV08372	Eucalyptus grandis
30	1201.8	33.1	3747	21	AA67145	Soybean cellulose
31	1183.4	32.6	2890	21	AA258270	Arabidopsis thalia
32	1177	32.5	3444	22	AAC83798	Cellulose synthase
33	1176.6	32.4	3828	19	AAV06566	Arabidopsis cellu
34	1147.2	31.6	3311	19	AAV08373	Cellulose synthase
35	1133.6	31.3	3232	22	AAC65448	Populus tremuloide
36	984.2	27.1	2248	19	AAV0562	Arabidopsis cellu
37	966.8	26.7	2125	21	AA258269	Soybean cellulose
38	933.6	25.7	2033	19	AAV08381	Cellulose synthase
39	921	25.4	2306	19	AAV34433	Cotton cellulose s
40	863.4	23.8	1401	21	AA44966	Arabidopsis thalia
41	835.4	23.0	2055	21	AA67111	Eucalyptus grandis
42	723.4	20.0	1860	21	AA67081	Eucalyptus grandis
43	666.2	18.4	1407	21	AA67086	Eucalyptus grandis
44	666.2	18.4	7234	22	AAC83247	Cellulose synthase
45	568.8	15.7	1741	19	AAV06569	Rice cellulose syn

ALIGNMENTS

RESULT 1
AA16458
ID AAS16458 standard; cDNA; 3799 BP.

XX AC AAS16458;

XX AC AAS16458;

DT 14-FEB-2002 (first entry)
DE Corn cDNA encoding cellulose synthase Cgrae19/cesA-19.

XX XX Corn; ss: cellulose synthase; Cdpqs45; cesA-3; Cgrae19; cesA-9;
KW stalk quality; improved stand; silage; pericarp; kernel hardening;
KW handling ability; transgenic plant.

XX OS Zea mays.

XX FH Key Location/Qualifiers

FT CDS 238..3799

FT /*tag= a

FT /*product= "Cellulose synthase"

XX PN WC200179516-A2.

XX PD 25-OCT-2001.

XX PF 12-APR-2001; 2001WO-US11951.

XX PR DNA encoding a mai

XX PR DNA encoding a mai

XX PR Wheat cellulose sy

XX PA Arabidopsis thalia

XX PI (PION-) PIONEER HI-BRED INT INC.

XX PI Dhugga KS, Helentjaris TG;

XX WPI; 2002-041338/05.
DR P-PSDB; AAU010496.
XX
PT New cellulose synthase polypeptides and polynucleotides, useful in
PT improving stalk quality or silage, and in increasing concentration of
PT cellulose in the pericarp, hardening the kernel for improved handling
PT ability.
XX
XX Claim 4; Page 80-85; 88pp; English.
XX
CC The invention relates to isolated nucleic acids encoding two cellulose
CC synthase proteins from corn, Cdp9545 (cesa-3) and Cqrae19 (cesa-9).
CC Also disclosed are a recombinant expression cassette comprising the
CC polynucleotide (operably linked to a promoter) a host cell comprising
CC the recombinant expression cassette and a transgenic plant comprising
CC the recombinant expression cassette. The nucleic acid is useful in
CC the improvement of stalk quality for improved stand or silage, and in
CC the increased concentration of cellulose in the pericarp, hardening
CC the kernel, and thus improving its handling ability. The nucleic acids
CC may also be used as probes or amplification primers in the detection,
CC quantification or isolation of gene transcripts, as probes in detecting
CC deficiencies in the level of mRNA, for detecting gene mutations or
CC allelic variants, for monitoring up regulation of expression or changes
CC in enzyme activity in screening assays, for site directed mutagenesis,
CC and in sense or antisense suppression of one or more genes in a host
CC cell, tissue or plant. The polypeptides may be used in assays for
CC enzyme agonists or antagonists, and as immunogen or antigen to
CC obtain antibodies specifically immunoreactive with the protein.
CC The present sequence encodes a corn cellulose synthase of the
CC invention.
XX
XX Sequence 3799 BP; 925 A; 931 C; 995 G; 944 T; 4 other;
XX
Query Match 71.68; Score 2595; DB 24; Length 3799;
Best Local Similarity 87.18; Pred. No. 0;
Matches 28/1; Conservative 0; Mismatches 420; Indels 4; Gaps 2;
XX
QY 36 CGGGTTGGATCGCTCGCGGCCATGGACGGCGGACGCCCTGAAAGTCGGGAGG 95
DB 214 CGGTTCTGGCGAGCTCGCCCTGCCATGGAGGGCGCGCGGGTGAAGTCGGGAGG 273
QY 96 CAGCGGGCGGGAGCGTGTGCACATCTGCGCGGCGGCGTGGCCACACCTGTGACGGC 155
DB 274 CGCGGGGAGGCGAGGTGTGCACATCTGCGGCGATGGCGCACATACGGCGAGGGA 333
QY 156 GAGCTCTTCACCGCTCGGACGTCTGCGCTCCCGCTGCGCCCTGCTAGGACAC 215
DB 334 GAGCTCTTCACCGCTCGGACGTCTGCGGTTCCCGGTGTGCGCCCTGCTAGGATAC 393
QY 216 GAGCGCAGGAGGGACCCAGCGCTGCTCCAGTGCAGACCAAGTACAGGGCCACAGA 275
DB 394 GAGCGCAGGAGGGACACAAAGCTGCCCGCTGCAAAACAAAGTACAGGGCCACAG 453
QY 276 GGGAGCCACAGCGATCCGCGGGAGGAGGCGGACGACACTGATCGCGATGATGATGAC 335
DB 454 GGGAGTCCAGCGATCCGAGGGGAGGAGGAGGAGGACTGATGCGCGATGATGATGAC 513
QY 336 TTCAACTACCTCGCATCTGGCAGTACAGGACCAAGCAAGATGCTGACAGGATGCGC 395
DB 514 TTCAACTACCTCGCATCTGGCAATGACGACCAAGCAAGATGCTGACAGGATGCGC 573
QY 396 AGCTGCCGATGAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGC 455
DB 574 AGCTGGCGCAWGAATGCTGGGGGAGCGGGGATGTTGGGGCGCCCAAGTATGACAGTGGT 633
QY 456 GAGATCGGCTCTCCAAGTATGACAGTGGAGATCCCTAGGGGATAGCTCCCTTCAGTC 515
DB 634 GAGATCGGCTTTACCAAGTACGACAGTGGTGGATCCCTCGGGGATACATCCCGTCAGTC 693
QY 516 ACCAAGCAGCGATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCATCATGATGCC 575
DB 694 ACTACAGCGAGATTTCGGGAGAAATCCCTGGTGTCTCCCTGACCATCATATGTCT 753

QY 576 CCTACGGGGAAATCATCAGCAGAC
DB 754 CCTACTGGGAACATTTGGCAGGC
QY 636 TCAAGGAGATTCTCCGGCAGTA
DB 814 TCGAGGGAATCTCTGTAGCG
QY 696 AAATGAAGCAGCAGCAGGTCG
DB 874 AAATGAAGCAGCAGCAGGGAAG
QY 756 GAAGGTCGGCAGCTACTGACAT
DB 934 GAGGGCGGGGTGTTGGTGATAT
QY 816 CTGAATGATCAAACTCGCCAGCG
DB 994 TTAACGATGAAACTCGCCAGCG
QY 876 AATCCCTACAGATGTCATTGT
DB 1054 AATCCATACAGGATGTCATTGT
QY 936 CGTCTCAAACTCCTGCGTAAG
DB 1114 CGGATCAAACTCCTGCGTAAG
QY 996 ATTTGGTTTCTTATCTCTGGATA
DB 1174 ATCTGGTTTCTCTTCTCTGGATA
QY 1056 GAGACCTACCTTGTATAGACTGCT
DB 1234 GAGACTTACCTTGTATAGACTGCA
QY 1116 GCTGCTGTGACATATTTGTCAATG
DB 1294 GCTGCTGTGACATATTTGTCAATG
QY 1176 GCCAACAACCTGCTATCCATTCTT
DB 1354 GCCAATACCTGCTATCCATTCTC
QY 1236 GTATCTGATCAGCAGGCTTCAATG
DB 1414 GTATCTGATCAGGAGCTGCTATG
QY 1296 GCTAGGAATGGGTACCATTTGTG
DB 1474 GCTAGAAAATGGGTGCGCATTTGT
QY 1356 TACTTTTCCAGAAAATTTGATTACG
DB 1534 TACTTCTCCAGAAAATTTGATTACT
QY 1416 CGCGGGCCCATGAAGAGAAATATG
DB 1594 CGCGGGCCCATGAAGAGAAATATG
QY 1476 AAGGCATTGAAAGTCCCGGAGGAAG
DB 1654 AAGGCACAAAAGTCCCTGAGGAAG
QY 1536 AACAAATACAGGATCATCTCGAA
DB 1714 AACAAATACAGGAGCATCTCGAA
QY 1596 GATACCTGAGGTAATCAGCTCCCGC
DB 1774 GATACCTGAGGTAATCAGCTACCCG

CCGTTTCCCTATGTGTAATCATTTCCACAAATCCG 635
CCATTTCCCTATATGTAATCATTTCAATCAATCCG 813
AATGTTCCCTGGAAAGAGAGAGATTGATGCTCG 695
AATGTTCCCTGGAAAGAGAGAGGTTGATGCTCG 873
CCCATGACTAATGGGACAGCATTTGCTCCCTCT 755
CCCATGACGAATGGCACAAGCATTTGCTCCCTCT 933
SCATCTACTGAATACACATGGAAGAGCGTTTA 815
SCATCAACTGATTACACATGGAAGATGCTTTA 993
CTAGAAAAGTCCCATTTGCTTCTTCCCAAAATA 875
CTAGAAAAGTTCACCTTCTCTCCAGGATA 1053
GGTTGGTTGTTCTAAGCATCTTCTCTGCACTAC 935
GATTGATTGTTCTAAGCATCTTCTTTCACACTAC 1113
ACCCACTGTGGCTTTTATCTGTATATGTAG 995
ACCCACTGTGGCTTCTATCTGTATATGTAG 1173
ATCAGTTCCGGAAGTGGTTTCCAAATCAACCG 1055
ATCAGTTTCCAAAGTGGTTTCCAAATCAACCG 1233
SGTATGACCGAGAGGTGAACGCTCTCAGTTG 1115
SGTATGACCGGGAAGGTGAACGCTCTCAGTTG 1293
TCGACCCCTTGAAGGAGCCACTATCTGCTACT 1175
TCGACCCCAATGAAGGAGCCCTCTCTGCTACT 1353
TGATATATCCGTTGACACAGGCTCTCTGCTAT 1235
GGACTATCTGTGGATAGGTTCTCTGCTAT 1413
TTTGTAGCAGCTTGGCTGAGACTTCCAGAGTTT 1295
TTTGTAGCAGCTTGGCTGAGACTTCCAGAGTTT 1473
STATGACATTTGAACCCAGAGCTCCCGAGTTT 1355
STACAACATTTGAACCTAGAGCTCTCTGAATGG 1533
AGACAAAGTCCAGCCCTTCATTTGTTAAAGAC 1415
AGACAAAGTCCAGCCCTTCATTTGTTAAAGAC 1593
ATTTAAATCAGGATAAATGCTTGTCTTCT 1475
ATCAAAATAGGTAATGCTTGTGCT 1653
ATCATGCAAGATGGCAGCACCATTGCCAGGA 1535
ATCATGCAAGATGGCAGCACCATTGCCAGGA 1713
CAGGTTTCTTGGTTCACAGTGGTGGCTT 1595
CAGGTTTCTTGGTTCACAGTGGTGGCTT 1773
GTTTATGTTCTGTCGTGTAAGAAAGCTCTGGG 1655
GTTTATGTTCTGTCGTGTAAGAAAGCTCTGGG 1833

1656 TTCCAGCACCACAGAGAGCGTGTGCCATGAATGCCCTTGTGCTGCTCAGCTGTCTT 1715
1834 TTCCAGCATCACAGAAAGTGTGTGCCATGAATGCTTGTGTCGCGCTCTCAGCTGTGTT 1893
1716 ACTAATGGACAATACATGTTGAATCTTGAATTTGATGATCACTACATCAACACAGCAGGCT 1775
1894 ACCAATGGACAATACATGTTGAATCTTGAATTTGATGATCACTACATCAACACAGTAAAGCT 1953
1776 GTCCGAGAAGCTATGTGTTCCPAATCGATCCAAACCTAGGTCGCCAAGTCTGTATGTG 1835
1954 CTCAGGGAAGCTATGTGTTCCCTTATGGATCCCTAACTAGGAAGGAGTGTCTGCTATGTT 2013
1836 CAGTTTCCACAAAGGTTTGAATGATGATAGGAATGATGATATCAACACAGGACACT 1895
2014 CAGTTTCCCGAGGTTCCGATGATTTGATAGGAATGATGATATGCCAACAGGAACACC 2073
1896 GTCCTTTTGTATTAAGTTCAGGGGCTTGCAGCGCATCAAGGACAGTTTATGTGGA 1955
2074 GTGTTTTCGATTAATCTTGAGAGGCTTGTATGGCATCCANGGACAGTTTATGTGGC 2133
1956 ACTGTTGTGTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAGGCGAAG 2015
2134 ACTGGCTGTGTTTCAACAGAACAGCTCTATATGTTTATGAGCCCCCAATTAAGCAAAAG 2193
2016 AAGCCAGTTTCTTGGCATCACTATGTGGGCAAGAGCAAGCAAGCAAGTCAAGAA 2075
2194 AAGGTGGTGTCTTGTCACTATGTGGTGGCAGGAAGGAAGCAAACTCAAGAA 2252
2076 AGGAGCTCAGATTAAGAAAGTGAACAAGCATGTGGACAGTCTGCTTCCAGTATTCAAT 2135
2253 --GGGCTCAGACAAGAAAGTCAAGAACGATGTGGACAGTCTGCTGCGAGTATTCAAT 2310
2136 CTCGAGACATAGAGAGGCTTGAAGTCTGAGTCTGGTTCATGATGAGAATCAGTTCTC 2195
2311 CTTGAAGATATAGAGGAGGAGTTGAAGGCGCTGGATTTGATGAGAATCAGTCTT 2370
2196 ATGCTCAAAATGAGCTTAGAGAAGAGATTGGCCAGTCAGCAGCATTTGTTGCCCTCACT 2255
2371 ATGCTCAAAATGAGCTTGGAGAAGAGATTGGCCAACTCTCAGGCTTTGTTGGCTCCACT 2430
2256 CTGATGGATATAGTGGTGTCTCAGTCTCCACTCCAGATCTCTTTGAAGAGCT 2315
2431 CTGATGGATATAGTGGTGTCTCAGTCTCCAGTCTCCAGTCTCCAGATCTCTCTGGAAGAGCT 2490
2316 ATCCATGTCTAAGTTGGGTATGAGGACAAAGTCTGAATGGGGAAGTGAATTTGGTGG 2375
2491 ATCCATGTCTAAGTTGGGTACGAGGACAAAGATTGAATGGGGAAGTGAATTTGGTGG 2550
2376 ATCTATGGATCTGTCACAGAGATATCTTACTGATTCAGATTCAGATCCAGCAGGCTGG 2435
2551 ATCTATGGTCTGTGACGGAAGATATCTCACTGGGTTCAGATTCAGATTCAGCAGGCTGG 2610
2436 CGTTCAGTCTATTGATGCCAAGCGCCAGCTTTCAAGGATCTGCCCTCCATCAATCTT 2495
2611 CGGTGATCTACTGATGCCTAAGCGCGCGCTTCAAGGATCGGCTCCCATCAATCTC 2670
2496 TCAGATCTGTCGAACAGATGCTGGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGC 2555
2671 TCAGACCGTCTGAACACAGTGTCCGGTGGGCTCTCGGTTCTGAGTGAATCTTTTCAGC 2730
2556 CGGATTTGCCCTTATGTTATGGCTACGAGGCGCTCTCAAGTTCCTGGAGAGATTCGCT 2615
2731 CGGCATTTGCCCTTATGTTATGGCTACGAGGCGCTCTCAAGTTCCTGGAGAGATTCGCC 2790
2616 TACATCAACACCACTATTAACCACTAACCTCTCTCCGCTTCTAGTCTATTTGATATG 2675
2791 TACATCAACACCACTATTAACCTCTACCTCTCCGCTCTCCCTCTCTCTATCTATCTCT 2850
2676 CTGCTCTCTGCTCTGATGGAAGTTTCATGTCAGAGATTCAGCACTTGGCCAGT 2735
2851 CTGTCATCTGCTCTCACGGGGAAGTTTCATGTCAGAGATTCAGCACTTCCGCTAGT 2910
2736 ATCTGTTTCATTCGCT 2795

2911 ATCTGTTTCATCT 2970
2796 AGTGGTGTGGCATTTGACAGAGTGTGGAGGAATGAACAGTTCCTGGTTCATTTGAGGTATG 2855
2971 AGCGGCTGGCATGACAGAGTGTGGAGGAACGAGCAGTTCCTGGTTCATTCGGGAGCATC 3030
2856 TCTGCAATCTCTGTTTGGCTCTTCTCAGGCTCTTCTGAAGGTGCTTGGCGGTATTCGACACC 2915
3031 TCGGCCCACTCTTTCGCGCTCTTCCAGGCGCTCTCAGGTGCTTGGCGGCATCGACACC 3090
2916 AACTTCACCTGTCACTCAAGGCTAATGACGAAGGCGACTTTGCTGAGTCTCTACATG 2975
3091 AACTTCACGCTCACCTCCAGGCTTCGGATGAAGACGCGGACTTCGCGGAGCTGTACATG 3150
2976 TTCAAGTGACAGCGCTTCTCATCCCTCCGACGACATTTTGTGATCATTAACATGTTGGT 3035
3151 TTCAAGTGACAGCACTTCTGATCCCGCCACCACTCTGATCATCAACCTGGTCGGC 3210
3036 GTCGTTGCTGGCACCTCTTACGCCATCAACAGTGTGTACCAATCATGGGGCGCTCTTT 3095
3211 GTTGTTCGCGGCACTCTTACGCCATCAACAGCGGTTACCACTGCTGGGTTCGCTCTTC 3270
3096 GGGAAGCTCTTCTTTCGCTTCTGGGTGATTTGTCACTTATATACCACTTCTCAAGGCTCT 3155
3271 GGCAGCTCTTCTTTCGCTTCTGGGTGATTTGTCACTTACCTGTACCCGTTCTCAAGGCTCT 3330
3156 ATGGGAGGCAAAACCGACACGAGGATTTGTCATGCTCTGGGCTGTCTCTCGCTCTCT 3215
3331 ATGGGTGCGCAGAACCGACCCGACCATCTGTTGCTGGGCGATCTCTGGGCTGCG 3390
3216 ATCTTCTCTTCTGCTGGGTGCTGTTGATCCATTCATCTACCTACCCGCTCTCGCTGGCCCAAT 3275
3391 ATCTTCTCTTCTGCTGGGTGCTGATCGATCGTTCACCAACCGCTCTCTGGCCGAT 3450
3276 ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGGA-GTTTGTAGAGACAGAA 3329
3451 ACTGNAACGTGTGGCATCAACTGCTAGGAGGTGGAAGGTTTGTAGAAACAGAGA 3505

RESULT 2
AAZ99512
ID AAZ99512 standard; DNA; 3746 BP.
XX
AC AAZ99512;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 321..3548
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1800..1802, aa: Xaa)
FT /note= "no termination codon given; Xaa is an
FT unspecified amino acid"
XX
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX

QY 1624 TAGTTATGTGTCGTGAAAGCGTCCTGGGTTCCAGCACCAAGAAGGCTGGTGCCA 1683
Db 1879 TGGCTATGTTTCTCGTGAAGGCGTCTGGATTCCAGCATCAACAAGAAAGCTGGTGCCA 1938
QY 1684 TCAATGCCCTGTGTCGTGTCACGTGTCCTACTAATGGACAATACATGTTGAATCTTG 1743
Db 1939 TGAATGCTCTGTTGCGTGTCTGAGCTGTGCTTACCAATGGACAATACATGTTGAATCTTG 1998
QY 1744 ATTGTGATCATTACATCAACACAGCAAGGCTGTCGCGAAGCTATGTGCTTCCCTAATGG 1803
Db 1999 ATTGTGATCATTACATCAACACAGCAAGGCTCTCAGGGAAGCTATGTGCTTCCCTAATGG 2058
QY 1804 ATCCAAACCTAGGTCGCGAAGTCTGTTATGTGCAGTTCGCCACAAGGTTTGATGGGATG 1863
Db 2059 ACCCTAAACCTAGGAAGAGTGTGCTGTACGTCCAGTTCCTCCAGAGATTCGATGGCAATG 2118
QY 1864 ATAGGAATGATCGATATGCAACACAGCAAGCTGTCTTTTGTGATTAACCTTGAGGGGCC 1923
Db 2119 ACAGGAATGATCGATATGCAACACAGCAAGCTGTGTTTTCGATATTAACCTTGAGAGTCT 2178
QY 1924 TTGACGGCATTCAGGACAGCTTTATGTGGAACTGTTGTTTTCGAAGAACAGCTA 1983
Db 2179 TTGATGGCATTCAGGACAGCTTTATGTGGAACCTGGCTGTGTTTTCGAACGGAACAGCTC 2238
QY 1984 TCTATGGTTATGAGCCCCCAATTAAGGCGAAGAACCCAGGTTTCTTGGCATCACTATGTG 2043
Db 2239 TATATGGTTATGAGCCCCCAATTAAGCAGAAGAGGGTGGTTCTTGTGATCACTATGTG 2298
QY 2044 GGGCAAGAAGAAGGCAAGCAAGTCAAGAAAGAGGCTCAGATGAAGAAAGTCGAACA 2103
Db 2299 GCGGTAGGAAGAAGGCAAGCAATCAAGAA---GGGCTCGGACAAGAAGTCGCAGA 2355
QY 2104 AGCATGTGACAGTTCGTGTCAGATTTCAATCTCGAAGACATAGAGAGGTTGTTGAAG 2163
Db 2356 AGCATGTGACAGTTCGTGTCAGATTTCAACCTTGAAGATATAGAGAGGAGTTGAAG 2415
QY 2164 GTGCTGGTTGATGATGAGAAATCAGTTCTCATGCTCAATGAGCTTAGAGAAGAGAT 2223
Db 2416 GCGCTGGATTGACGACGAGAATCACTTCTTATGTCTCAATGAGCCTTGAGAAGAGAT 2475
QY 2224 TTGGCAGTCAGCAGCATTTGTCCTCCACCTGATGGAATATGGTGTGTTCTCAGT 2283
Db 2476 TTGGCAGTCGCGAGGTTGTTGCTCCACTGATGAGATATGGTGTGTTCTCAGT 2535
QY 2284 CTTCCATCCGAATCTTTTGAAGAAGCTATCCATGTCATAAGTTGTGCTATGAGG 2343
Db 2536 CCGCAACTCCGGAGTCTCTTCTGAAGAAGCTATCCATGTTATAAGCTGTGGCTATGAGG 2595
QY 2344 ACAAGTCTGAATGGGAACAGATGTTGGATCTATGGATCTGCACAGAAGATATTC 2403
Db 2596 ACAAGACTGAATGGGAACAGATGCGGTGGATCTACGGTCTGTGACAGAAGACATTC 2655
QY 2404 TTACTGATTCAGATGTCACGCAAGAGGCTGGCGTTCAAGTCTATTTGATGCCCAAGCGCC 2463
Db 2656 TCACCGGATTCAGATGTCACGCGAGGCTGGCGTTCGATCTACTGCAATGCCCAAGCGCG 2715
QY 2464 CAGCTTCAAGGATCTGCCCAATCAATCTTTCAGATCTGCTGAACCAAGTCTGCGGT 2523
Db 2716 CAGCTTCAAGGATCTGCCCAATCAATCTTTCGACCTCTGAACCAAGTCTGCGGT 2775
QY 2524 GGCTCTCGGTTCTGTTGAATTTCTTTTCAGCGGCGATTGCCCCCTTATGGTATGGCTAGG 2583
Db 2776 GGCTCTGTTGGTCCGTTGAGATCTCTTTCAGCGGCGACTGCCCTGTGTAGCGCTAGG 2835
QY 2584 GAGGGCGCTCAAGTTCCTGGAGAGATTCGTTTACATCAACACCAACCATTTACCACTAA 2643
Db 2836 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCAACCATTTACCGGCTCA 2895
QY 2644 CTTCTCCGCTCTAGTCTATTGTATATGCTCTCTATGCTGCTCACTGGAAGT 2703
Db 2896 CGTCCATCCGCTCTCTCATCTACTGTCATCTGCCGCCCATCTGCTGCTACCGGAAGT 2955

QY 2704 TCATCATGCCAGAGATTAGCAACTTGGCCAGATATCTGGTTCATTTGGCTCTTCCTTCAA 2763
Db 2956 TCATCATTCAGAGATCAGCAACTTTCGCCAGCATCTGGTTCATCTCCCTCTTTCATCTCGA 3015
QY 2764 TTTTCCCCCACTGTGATTCCTTGAGATGAGTGGAGTGGTGTGGCATTTGACGAGTGGTGA 2823
Db 3016 TCTTCCACAGCGGCATCTCGAGATGAGTGGAGCGGGTGGCATCGACGAGTGGTGA 3075
QY 2824 GGAATGAACAGTTCTGGGTCATTTGGAGGTATCTTCACATCTGTTTGGCGTCTTCAGG 2883
Db 3076 GGAACAGAGAGTTCTGGGTGATCGGGGCATCTCCGCGACCTCTTCGCGGTGTTCCAGG 3135
QY 2884 GTCTTCTGAAGTGTTCGCGGTATCGACACCAACTTCACTGTCACTCAAAAGGCTAATG 2943
Db 3136 GCTGTCAAGTGTTCGCGGCATCGACACCAACTTCACTCACTCAAGGCTCGG 3195
QY 2944 ACGAAGAAGCGACTTGTGTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCCCTC 3003
Db 3196 ACGAGGACGCGACTTTCGCGGAGCTGTACATGTTCAAGTGGAGCGCTCTGATCCCGC 3255
QY 3004 CGAGGACCATTTTGATCATTAACATGTTGGTGTCTGCTGTCACCTCTAGGCCATCA 3063
Db 3256 CCACCCACCATCTGATCAATCACTGTCGCGGTCTGCGCGCATCTCTAGCCCATCA 3315
QY 3064 ACAGTGTGTTACCAATCATGCGGCGCGCTCTTTGGGAAGCTCTTCTTTGGCTTCTGGGTGA 3123
Db 3316 ACAGCGATACCAAGTGTGGGCGCGCTCTTCGGAAGCTCTTCTTTCGCTTCTGGGTCA 3375
QY 3124 TTGTTCACTTATACCAATTCCTCAAGGTTCTTATGGGAGGCAAAACCGCACACCGACGA 3183
Db 3376 TCGTCCACCTGTACCGGCTCTCAAGGCGCTCATGGGCGCATGGGCAAGAACCCCGACCA 3435
QY 3184 TTGTCATCGTCTGGCTGTCCTCTCGTTCATCTTCTCTGCTGCTGGTTCGTGTTG 3243
Db 3436 TCGTGTGCTGTGGCGCATCTCTGTCGCTTCATCTTCTTCTGCTGGGTTTCGATCG 3495
QY 3244 ATCCATTCACTACCGCTCTCGCTGGCCCAATATCCAAACCTTGGCATCAACTGCTAGG 3303
Db 3496 ACCCTTTCACCAACCGGCTCACTGGCCGGATACCCAGACGTGTGCACTCAACTGCTAGG 3555
QY 3304 AAAGTGGAGTTT 3316
Db 3556 GAAGTGAAGGTT 3568

RESULT 3
AAZ99527 standard; DNA; 3746 BP.
XX
AC AAZ99527;
XX
DT 03-JUL-2000 (first entry)
XX

DNA encoding a maize cellulose synthase.

Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
transgenic plant; plant breeding marker; ss.

Zea mays.

Key Location/Qualifiers
CDS 321..3449
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 1800..1802, aa: Xaa)
FT /note= "no termination codon given; Xaa is an
unspecified amino acid"

WO200009706-A2.

24-FEB-2000.

16-AUG-1999; 99WO-US18760.

XX PR 17-AUG-1998; 98US-0096822.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX XX WPI; 2000-224343/19.
DR DR P-PSDB; AAY84119.
XX XX
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX XX
PS Claim 1; Page 176-181; 119pp; English.
XX The present sequence encodes a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX XX
SQ Sequence 3746 BP; 861 A; 976 C; 1006 G; 902 T; 1 other;
Query Match 70.9%; Score 2572.2; DB 21; Length 3746;
Best Local Similarity 86.4%; Pred. No. 0;
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;
QY 4 CGAGNACCCGCTCCAGCTCTGTCGTCGTCGGGTGGATCGCTCTGCGCGCCCATGG 63
DB 265 CCAGGCCCCAGGCTCCAGGCCAGCTCCCTCGACGCTTTCTGGCGAGCTCGCTTGCCTATGG 324
QY 64 ACGGGGACGCGGACGCGCTGAAGTCCGGGAGGCACGGGCGGCGAGCTGTGCCAGATCT 123
DB 325 AGGCGGACGCGGACGCGCTGAAGTCCGGGAGGCAGCGGTGGCGACAGGTGTGCCAGATCT 384
QY 124 GCGCGGACGCGCTGGGACCGAGTGGAGCGGACGCTCTTACCGCCTGCGAGCTGTGCC 183
DB 385 GCGGCGACGCGCTGGGACCGAGCGGAGGCGGACGCTTTCGCGCGCTTGGGAGCTGTGG 444
QY 184 GCTTCCCGTCTGCGCCCTGCTACGAGCAGCAGCGGAGGCGGAGGCGGAGGCGGCTGCC 243
DB 445 GTTTCCCGTGTGCGCCCTGCTACGAGTACGAGCGGAGGCGGAGGCGGAGGCGGCTGCC 504
QY 244 TCCAGTGAAGACCAAGTACAAGCGCCACAGGAGGAGGCGGAGGCGGAGGCGGAGGAG 303
DB 505 CCAGTGAAGACCAAGTACAAGCGCCACAGGAGGAGGCGGAGGCGGAGGCGGAGGAG 564
QY 304 GCGAGCAGACTGATCCCATGATGGTAGTGACTTCACTACCTGATCTGGCAGCTGAGG 363
DB 565 GAGAGCAGACTGATGCCGA-----TAGCGACTTCAATACCTTGCATCTGGCAATAGG 618
QY 364 ACCAGAAGCAGAAGATTGCTGACAGGATGCGCAGCTGGCGCATGAACACCGGGGCGAGT 423
DB 619 ACCAGAAGCAGAAGATTGCGGACAGATGCGCAGCTGGCGCATGAACGTTGGGGCAGCG 678
QY 424 GCANTGTGGCCACCCCAAGTATGACAGTGGCGAGATTCGGCCTCTCCAAAGTATGACAGTG 483

DB 679 GGGATGTTGGTCGCGCCCAAGTATGACAGATGCGGCGCTTACCAAGTATGACAGATG 738
QY 484 GAGAGATCCCTAGGGGATACGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 543
DB 739 GCGAGATTCCTCGGGGATACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 798
QY 544 CTGGAGCTTCGCTGATCATCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 603
DB 799 CTGGTGTCTCCCTGACCATCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 858
QY 604 GCTTTCCTATGTGAATCATTCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 663
DB 859 CATTTCCCTATGTGAACCATTCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 918
QY 664 ATGTTGCTCGGAAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 723
DB 919 ATGTTGCTCGGAAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 978
QY 724 CCATGACTAATGGGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 783
DB 979 CCATGACTAATGGGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1038
QY 784 CATCTACTGAATACAAACATGGAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 843
DB 1039 CATCAACTGATTACAACATGGAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1098
QY 844 CTAGAAAGTCCCATGCTTCGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 903
DB 1099 CTAGAAAGTCCCATGCTTCGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1158
QY 904 GGTGTTGTTCTAAGCATCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 963
DB 1159 GATTGATTGTTCTAAGCATCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1218
QY 964 ACCCACTGCGCTTCTATCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1023
DB 1219 ACCCACTGCGCTTCTATCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1278
QY 1024 ATCAGTTCCTCGAAGTGGTTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1083
DB 1279 ATCAGTTCCTCGAAGTGGTTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1338
QY 1084 GGTATGACCGAGAGAGTGAACCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1143
DB 1339 GGTATGACCGAGAGAGTGAACCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1398
QY 1144 TCGACCCCTTGAAGGAGGACCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1458
DB 1399 TCGACCCCAATGAAGGAGGCTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1518
QY 1204 TTGATTATCCCGTGGACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1263
DB 1459 TGGATTACCCCTGTGGATGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1518
QY 1264 CTTTTCGACGATGCTGCTGAGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1323
DB 1519 CATTTGATGCACTAGCTAGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1378
QY 1324 AGTATGACATTGAACCCAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1383
DB 1579 AGTAAACATGAACCTAGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1638
QY 1384 AGACAAAGTCCAGCTTCATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1443
DB 1639 AGGACAAAGTCCAGCTTCATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1498
QY 1444 AATTTAAAGTGAAGTCCCGAGGAGGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1503
DB 1699 AATTCAAAGTGAAGGTAATGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1758
QY 1504 GGATCATGCAAGATGGCACCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1563

1716 AATTCAAAGTTAGGGTAAATGGCCCTTGTGCTAAGGCACAGAAAGTTCTGAGGAAGAT 1775
1504 GGATCATCAAGATGGGCACACCATGGCCAGGAACAAATACAGGGATCATCTCTGGAATGA 1563
1776 GGATCATCAAGATGGGCACACCATGGCCAGGAACAAATACCGGGACCAATCTCGAATGA 1835
1564 TTCAGGTTTCTCTGGTGCACAGTGGGCGCTTGATATCTGAGGGTAAAGAGTCCGCCGTT 1623
1836 TTCAGGTTTCTCTGGTGCACAGTGGGCGCTTGATATCTGAGGGCAATGAGCTACCCCGTT 1895
1624 TAGTTTATGTCTCGTGAAGCGTCTTGGGTTCCAGACACACAGAAAGCGTGGTGCCA 1683
1896 TGGTCTATGTTCTCGTGAAGCGTCTTGGATTCCAGCATCACAGAAAGCTGGTGCCA 1955
1684 TGAATGCCCTTGTCTGGTCTCAGCTGCTCTTACTAATGGACAATACATGTTGAATCTTG 1743
1956 TGAATGCTCTTGTCTGGTCTCAGCTGCTTACCAATGGACAATACATGTTGAATCTTG 2015
1744 ATTGTGATCACTACATCAACAACAGCAAGGCTGTCCGAGAAGCTATGTCTTCCCTAAATGG 1803
2016 ATTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGGAAGCTATGTCTTCCCTATGG 2075
1804 ATCCAAACCTAGTCCGCAAGTCTGTATGTGCAGTTCGCCAAGGTTTGATGGGATG 1863
2076 ACCCTAAGCTAGGAAGGAGTGTCTGTACGTCCAGTTCGCCAGATTCGATGGCATTG 2135
1864 ATAGGAATGATCGATATGCAAAACAGGAACACTGTCTTTTGTATATTAATCTTGAAGGGCC 1923
2136 ACAGGNATGATGATATGCCAACAGGAACACCGTGTTTTCGATATTAATCTTGAGAGTTC 2195
1924 TTGACGGCATTCAGGACCACTTATGTGGAACTGGTGTGTTTTCACACAAGCACTA 1983
2196 TTGATGGCATCAAGGACCACTTATGTGGAACTGGTGTGTTTTCACACAAGCACTA 2255
1984 TCTATGGTTATGAGCCCAATTAAGCGAAGAGCCAGGTTTCTTGGCATCACTATG 2043
2256 TATATGGTTATGAGCCCAATTAAGCAGAAGAGGTTGTTTCTGTCATCACTATG 2315
2044 GGGCAAGNAGNAGCAGCAAGTCAAGAAAGAGGCTCAGATAGAAAGATCGACACA 2103
2316 GGGGTAGGAAGNAGCAAGCAAAATCAAGAA ---GGGCTCGGCAAGAAAGATCGCAGA 2372
2104 AGCATGTGACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGGTGTGAAG 2163
2373 AGCATGTGACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGGAGTTGAAG 2432
2164 GTGCTGGTTGATGATGAGAATCAGTCTCATGTCTCAATAGCTTAGAGAGAT 2223
2433 GGCCTGGATTGACGACGAGAATCACTTCTATGTCTCAATAGCCTGGAGAGAGAT 2492
2224 TTGGCCAGTACAGCAATTTGTGCCCTCCACTGTGATGGAATATGGTGGTGTTCCTCACT 2283
2493 TTGGCCAGTCCAGCGTTTGTGCTCCACTGTGATGAGATGGTGGTGTTCCTCACT 2552
2284 CTTCCACTCCAGATCTTGTGAAGAAGCTATCCATGTGCATAGTTGGCTATGAGG 2343
2553 CCGCAACTCCGAGTCTCTTGAAGAAGCTATCCATGTATAGCTGGCTATGAGG 2612
2344 ACAAGTCTGAATGGGAACTGAGATGGTGGATCTACGGTTCGTGACAGAAGATATTC 2403
2613 ACAAGTCTGAATGGGAACTGAGATGGTGGATCTACGGTTCGTGACAGAAGATATTC 2672
2404 TTACTGGATTCAAGATGACCAAGAGGCTGGCGTTCAGTCTATTCATGCCCAAGCGCC 2463
2673 TCACCGGATTCGAAGATGACCGGAGGCTGGCGTGGATCTACTGATGCCCAAGCGGC 2732
2464 CAGCTTTCAGGATCTGCCCGCATCAATCTTTTCAGATCTCTGTAACCAAGTCTCCGGT 2523
2733 CAGCTTTCAGGATCTGCCCGCATCAATCTTTTCAGATCTCTGTAACCAAGTCTCCGGT 2792
2524 GGGCTCTCGGTTCTGTTGAATCTTTTACCGCGGCAATGCCCTTATGGTATGGCTAGG 2583

2793 GGGCTCTTGGTTCGTCGAGATCTCTTCAAGCGGCACTGCCCGCTGTGTAGGGCTAGC 2852
2584 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACACACCATTTACCCACTAA 2643
2853 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCGTACATCAACACACACCATCTACCCGCTCA 2912
2644 CTTCTCTCCCGTCTTCTAGTCTTATTTGATATTTGCTGCTATCTCTGCTACCTACCTGGAAGT 2703
2913 CGTCCATCCCGTCTTCTATCTACTGCTATCTGCGCCGCACTCTGTCTGCTCACCAGAAAT 2972
2704 TCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCAATTTGGCTCTTCTCTTCAA 2763
2973 TCATCATTCAGAGATCAGAACTTGCACAGATCTGGTTCATCTCCTCTTTCATCTCA 3032
2764 TTTTGGCACTGGTATCTTGGATGAGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2823
3033 TCTTCCCAAGGCACTCTGGAGATGAGTGGAGCGGGGTGGGATCGAGGATGGTGG 3092
2824 GGAATGAACAGTCTGGGTCATTTGGAGTATCTCTGCACATCTGTTTGGCGTCTTTCAGG 2883
3093 GGAACGAGCAGTCTTGGGTGATTCGGGGCATCTCCGGGCACCTCTTCCCGCTTTCAGG 3152
2884 GTCTTCTGAAGTCTTGCCTGATTCGACACCAACTTCACTGTCCACCTCAAGAGCTAATG 2943
3153 GCTGCTCAAGTCTTGGCGGCATTCGACACCAACTTCACTGTCCACCTCAAGAGCTCGG 3212
2944 ACGAAGAGGCACTTCTGCTGAGCTCTACATGTTCAAAGTGAGACAGCTTCTCATCCCTC 3003
3213 ACGAGACGGCACTTTCGGGAGCTGTACATGTTCAAAGTGAGACAGCTTCTCATCCCGC 3272
3004 CGACGACCAATTTGATCATTAACATGTTTGGTGTCTGCTGGCAGCTCTACGCGCATCA 3063
3273 CCACCACTCTGATCATCAACCTGCTGGGCTGCTGCGGCACTCTCTACGCCATCA 3332
3064 ACAGTGGTTACCAATCATGCGGCGCTCTTTGGGAAAGCTCTTCTTGGCTCTTGGGTGA 3123
3333 ACAGGATACAGCTGTTGGGCGGCTCTTCCGCAAGCTCTTCTTCCGCTTCTGGGTCA 3392
3124 TTCTTCACTTATACCATTCCTCAAGGCTCTTATGGGAGGCAAAACCCGACACACGACGA 3183
3393 TCGTCCACCTGTACCGTCTCTCAAGGCTCTCATGGGAGGCAAGACCGCACCCGACCA 3452
3184 TTGTCATCTGCGGCTGTCT 3243
3453 TCCTCTCTCTGCGCATCTCTGCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3512
3244 ATCCATTCATCTACCGCTCTCTCTGCGCCCAATATCCAAACCTGTGGCATCAACTGTAGG 3303
3513 ACCCTTCAACACCGCTCACTGCGCGGATACCAGAGCTGTGGCATCAACTGTAGG 3572
3304 AAAGTGGGAGTTT 3316
3573 GAAGTGAAGGTT 3585

RESULT 5

AA299533
ID AA299533 standard; DNA; 3704 BP.

XX AA299533;

XX AC
XX DT 03-JUL-2000 (first entry)

XX XX
XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers
XX FT 272..3496
XX CDS /*tag= a

FT /product= "cellulose synthase"
 FT /transl_except= (pos: 1544...1546, aa: Xaa)
 FT /note= "no termination codon given; Xaa is an
 XX unspecified amino acid"

PN WO200009706-A2.

PD 24-FEB-2000.

PF 16-AUG-1999; 99WO-US18760.

PR 17-AUG-1998; 98US-0096822.

PA (PION-) PIONEER HI-BRED INT INC.

PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI: 2000-224343/19.

DR P-PSDB: AA184121.

XX New genes which encode maize cellulose synthase polypeptides in plants
 PT useful for modulating the expression of cellulose synthase in plants
 PT and to produce transgenic plants expressing the novel protein

PS Claim 1; Page 191-196; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide.
 CC The cellulose synthase can be used for the improvement of stalk quality
 CC for improved stand or silage. It also provides an increased concentration
 CC of cellulose in the pericarp, hardening the kernel and improving its
 CC handling ability. The sequences are used to produce transgenic plants
 CC and seeds expressing the cellulose synthase. The polynucleotide is
 CC used for modulating, preferably increasing, the level of the synthase
 CC in a plant cell. The plants are preferably monocots. The polynucleotide
 CC is also used as a probe or primer in the detection quantitation or
 CC isolation of gene transcripts. The probes are useful in detecting
 CC deficiencies in the level of mRNA in screenings for desired transgenic
 CC plant, for detecting mutations in the gene, for monitoring upregulation
 CC of expression or changes in enzyme activity in screening assays of
 CC compounds, for detection of any number of allelic variants of the gene,
 CC or for use as molecular markers in plant breeding programs. The
 CC isolated nucleic acids of the present invention can also be used for
 CC recombinant expression of their encoded polypeptides or for use as
 CC immunogens in the preparation and/or screening of antibodies. The
 CC proteins can be employed in assays for enzyme agonists or antagonists
 CC of enzyme function or for use of immunogens or antigens to obtain
 CC antibodies specifically immunoreactive with a protein.

XX Sequence 3704 BP; 909 A; 838 C; 962 G; 994 T; 1 other;

50 Query Match 64.1%; Score 2325.2; DB 21; Length 3704;
 Best Local Similarity 82.8%; Pred. No. 0;
 Matches 2706; Conservative 0; Mismatches 549; Indels 12; Gaps 4;

Qy 38 GGTGGATCGCTCGCGCGCATGAGCGGCGACGCCCTGAAGTCCGGGAGCGA 97
 Db 250 GGAGGGCTCGGATGTGGTGGCCATGGACGGCGGCG---GACGCCACAGAAATTCGGGGGAAGCA 306
 Qy 98 CGGGCGGGGAGCGTGTGCAGATCTGGCGCGACGGCTGGCGCACCACTGTGGACGCGCA 157
 Db 307 TGTGGCGGGCAGGTGTGCAGATCTGGCGGACGGCTGGCGCACCGCGCGGACGCGCA 366
 Qy 158 CGTCTTCACCGCTCGGAGCTCGCGCTCCCGCTCGCGCGCGCTGCTACGAGCACGA 217
 Db 367 CCTCTTCACCGCTCGGAGCTCGCGCTCCCGCTGCGCGCGCTGCTACGAGTACGA 426
 Qy 218 GCGCAAGGAGGCGACCGAGCTGCTCCAGTGCAGAACCAAGTACAAAGCGCCACAGAG 277
 Db 427 GCGCAAGGAGGCGACCGAGCTGCTCCAGTGCAGAACCAAGTACAAAGCGCCACAAAG 486
 Qy 278 GAGCCAGGCGATCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 337
 Db 487 GAGCCAGGCGATCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546

Qy 338 CAACTACCCTGCATCTGGGCACT 397
 Db 547 CAACTACCAGCATCTGGCAAC 606
 Qy 398 CTGGCGCATGAACACCGGGGGC 457
 Db 607 TTGGCGCAAACTCACGTGGC 663
 Qy 458 GATCGGCTCTCCAAGTATGAC 517
 Db 664 AATGGGCATGGGAAGTATGAC 723
 Qy 518 CAACAGCCAGATGTCAGGAGAA 577
 Db 724 TCATAGCCAGATCTCAGGAGAG 783
 Qy 578 TAGGGGAACATCAGCAGACGTG 637
 Db 784 TGGGAACATTTGGCAGCGCTGGAC 843
 Qy 638 AAGGGAGTCTTCGGCAGTATTC 697
 Db 844 GAGGGAGTCTTCGGTAGCCTTG 903
 Qy 598 AATGAAGCAGGACAAGGTGGCA 757
 Db 904 AATGA---AGGATAAAGGTGCAR 960
 Qy 758 AGSTCGGGCAGCTACTGACATCG 817
 Db 961 AGGCGTGGAGTGTCTGATATTC 1020
 Qy 818 GAATGATGAACCTCCGCCACCTC 877
 Db 1021 GAATGATGAACCTCCGCCACCTC 1080
 Qy 878 TCCTCAGAAAGGTGCATTTGTT 937
 Db 1081 TCGGTACAGAAAGGTGCATTTG 1140
 Qy 938 TCTCACAAATCCCTGTCGTAATG 997
 Db 1141 TATCACACATCTCTGTAACAATG 1200
 Qy 998 TTGGTTGCTTTATTCCTGATAC 1057
 Db 1201 CTGGTTGCTTTGCTCCTGATTT 1260
 Qy 1058 GACTACCTTGATAGACTGGCTTT 1117
 Db 1261 AACATACCTGATAGACTGGCTTT 1320
 Qy 1118 TGCTTTGACATTTTGTTCAGTAC 1177
 Db 1321 TCCTGTTGATATTTTGTTCAGTAC 1380
 Qy 1178 CAACACTGTGCTATCCATTCTTGC 1237
 Db 1381 AAATACTGTGCTTCCATCCTTGC 1440
 Qy 1238 ATCTGATGACGGAGCTCAATGCT 1297
 Db 1441 TTCGGATGAGGAGCTGCTATGCT 1500
 Qy 1298 TAGAAATGGTACCAATTTGTGAA 1357
 Db 1501 TAGAAATGGTTCCTCTTCTGTA 1560
 Qy 1358 CTTTTCGACAGAAATGATTACCTC 1417
 Db 1561 CTTTGTGACAGAAATGATTACTTC 1620

Qy	1418	CCGGCCATGAGAGAGAAATATGAAGAAATTTAAATCAGGATAAATGGCCCTAGTTCTTCAA	1477
Db	1621	CCGGCCATGAAGAGAGAAATATGAAGAAATTTAAATCAGGATAAATGGCTGTGTAGCCCAA	1680
Qy	1478	GGCATTGAAGTCCCGAGGAGGATGATCATCAAGATGGCACACCATGGCCGACGAAA	1537
Db	1681	GGCAGAAAAGTTCCGAGGAGGATGATCATCAAGATGGTATACACCTTGGCTGGGAA	1740
Qy	1538	CAATACAGGGATCATCTCGGAATGATTCAGGTTTTCTTGGTCACAGTGGTGGCCTTGA	1597
Db	1741	CAATACTAGGACCATCTCGGAATGATTCAGGTTTTCTTGGTCACAGTGGAGGGCTTGA	1800
Qy	1598	TACTGAGGTAAATGAGCTCCCGCTTTAGTTTATGTGTCGTGAAAAGCGTCTCGGTT	1657
Db	1801	CGTTGAGCAATGAATCTCCCTGTTTGGTTATGTGTCGTGAAAAGCGTCTCGATT	1860
Qy	1658	CCAGCACCAAGAAGCGTGGCCATGAATGCCCTTGTTCGTGTCACAGCTGTCCTTAC	1717
Db	1861	CCACATCACAGAAGCGTGGCCATGAATGCACCTTGTTCGTGATCAGCTGTCCTTAC	1920
Qy	1718	TAATGGCAATACATGTTGAATCTTTGATGTGTGATCACTACATCAACAACAGCAAGCGT	1777
Db	1921	TAATGGCAATACATGTTGAATCTTTGATGTGTGATCACTACATCAATAATAGCAAGCTCT	1980
Qy	1778	CCGAGAAGCTATGTCCTTCCATAATGGATCCAAACCTAGGTCGCAAGCTCTGTTATGCA	1837
Db	1981	TCGAGAAGCTATGTCCTTCCATAATGGATCCAAACCTAGGATGCAAGCTCTGTTATGCA	2040
Qy	1838	GTTCCCAAAAGTTTATGGGATTTGATAGGAATGATCGATATGCAACAACAGCAACACTGT	1897
Db	2041	ATTTCTCAGAGTTTATGATGTTATGNTAGNAATGACCGATATGCAACAACAGCAACACTGT	2100
Qy	1898	CTTTTGTGATTAATCTGAGGGCCTTGACGGCAATCAAGGACCACTTTATGTGGGAAC	1957
Db	2101	GTTTTTCGATTAATCTGAGAGCTCTTGACGGCAATCAAGGGCCAGTTATGTGGGAAC	2160
Qy	1958	TGGTTGTGTTTTCAACAGAACAGCTATCTATGGTTATGAGCCCCCAATTAAGCGGAGAA	2017
Db	2161	TGSGTTGTGTTTTAAACAGAACCGCTTATATGGTTATGAGCTCCAGTCAAGAAAAAAA	2220
Qy	2018	GCCAGGTTCTTGCCATCACTATGTGGGGCAAGAAAGCAAGCAAGTCAAGAAAAAG	2077
Db	2221	GCCAGGTTCTTGCCATCACTATGTGGGGCAAGAAAGCAAGCAAGTCAAAATCTAAGAA	2277
Qy	2078	GAGCTCAGATAAGAAAAAGTCGAACAAGCATGTGACAGTTCCTGTTCAGATTAATCAATCT	2137
Db	2278	GAGCTCGAAAAAGAAAGTCAATAGACACGACGACAGTTCCTGTACCAAGTATTTAATCT	2337
Qy	2138	CGAAGACATAGAGAGGGGTTCGAAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCAT	2197
Db	2338	CGAAGATATAGAGAGGGGATTGAAGGTTCTCAGTTTTGTATGAGAAATCGCTGATTAT	2397
Qy	2198	GTCCTCAATGAGCTTAGAGAAGAGATTGGCCAGTCAGCAGATTGTTGCTCCACTCT	2257
Db	2398	GTCCTCAATGAGCTTAGAGAAGAGATTGGCCAGTCCAGTGTGTTGTAGCCTCTACTCT	2457
Qy	2258	GATGGAATATGGTGTTCCTCAGTCCTCCACTCCAGATCTCTTTTGAAGAAGCAT	2317
Db	2458	GATGGAATATGGTGTTCCTCAGATCTCCAACTCCAGAGTCTCTTCTGAAAGAAGCAT	2517
Qy	2318	CCATGTCATAGTTGTGGCTATGAGGACAAAGTCTGNTGGGGAACTGAGATTTGGTTGAT	2377
Db	2518	TCATGTCATCAGCTGTGGCTATGAGGACAAAACCTGACTGGGAACTGAGATTTGGTGGAT	2577
Qy	2378	CTATGGATCTGTACAGAAAGATTTCTTACTGGATTCRAGATGCACGCAAGGCTCGCG	2437
Db	2578	CTATGGTGTCTGTACAGAAAGCATCTCCACGGATTCRAGATGCATGCTCCAGGCTCGCG	2637
Qy	2438	TTTCACTTATGTCATGCCAAGCGCCAGCTTTTCRAGGGATCTGCCGCCCATCAATCTTTC	2497
Db	2638	ATCAATCTACTGTCATGCGCTAAGCGACACGTTTTCRAGGGATCTGCTCTCTATCAACCTTC	2697
Qy	2498	AGATCGTGTGAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGCGC	2557

Db	2698	GGATCGTTTGAATCAAGTGCCTTCGGTGGGCTCTTGGTTCCATTGAAATCTTTTCAGCAG	2755
QY	2558	GCATTCGCCCTTTATGGTATGCGTACGGAGGCGGCTCAAGTTCCTGGAGAGATTCGCTTA	2617
Db	2758	GCATTTGCCCATATGGTATGGCTATGGAGCGGCTTAAATTCCTGGAGAGATTTGCTTA	2817
QY	2618	CATCAACACCACTTTACCCACATACTCTCTCCGCTCTTAGTCTATTGTATATTGCC	2677
Db	2818	TATCAACAACAATTTATCCACTCACATCAATCCCGTCTCTCTGACTGCATATTGCC	2877
QY	2678	TGCTATCTCTGCTGCCTCACTGGAAAGTTCATCATGCCAGAGATTAGCAACTTGCCCAAGTAT	2737
Db	2878	AGCAGTTTGTCTTCACCTGGGAAGTTCATCATCCCAAGATTAGTAACCTAGAGAGTGT	2937
QY	2738	CTGGTTCAATTCGGCTCTTCCTTTCAATTTTCGCCACTGGTATCCTTCAGATGAGGTGGAG	2797
Db	2938	TTGGTTTATATGGCTCTTTATCTCAATCTTTTGCCACTGGTATCCTTCAGATGAGGTGGAG	2997
QY	2798	TGGTGTTCGANTGACGAGTGTGTGAGGAATGAACAGTTCCTGGGTCAATTCGGAGGTATCTC	2857
Db	2998	TGGTGTTCGCAATTTGATGAATGGTGGAGGACGACGATTCCTGGGTCAATTCGGTGTATTC	3057
QY	2858	TGCACATCTGTTTGGCGTCTTTTCAGGGTCTTCTGAAGGTGCTTCGCCGTATCGACACCAA	2917
Db	3058	TGCGCATTTATTTGCGGTCTTCCAGGGTCTCCTGGAAGTGTCTTCTGCTGTATCGACACGAG	3117
QY	2918	CTTCACTGTCACTCAAGGCTTAATGACGAGAGAGGCGACATTTGCTGCTGAGCTCTACATGTT	2977
Db	3118	CTTCACTGTCACTCTTAAGGCCACTGAGGAAGAGGTGATTTTGGCAGGCTCTACATGTT	3177
QY	2978	CAAGTGGACGAGCTTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGGTTGGTGT	3037
Db	3178	CAAGTGGACACAGCTTCTGATCCCAACCAACCACTATTTTGATCATTAACCTGGTCCGCGT	3237
QY	3038	CGTTGCTGGCACCTCTCAGCCCATCAACAGTGGTTTACCAATCATGGGGCGGCTCTTTGG	3097
Db	3238	GGTCTGTGGCATTTCTTACGCAATCAATAGCGGTTACCAGTCATGGGAGCCCTCTTTCCG	3297
QY	3098	GAAGCTCTCTTTGGCTCTTCGGGTGATTTGTTCACTTATACCGATTCCTCAAGGGTCTTAT	3157
Db	3298	GAAGCTCTCTTTGGCTCTTCGGGTGATTTGTCACCTGTACCCCTTCCTCAAGGGCCCTCAT	3357
QY	3158	GGCAGGCAAAACCGCACACCGACGATTTGTCATCTGCTGGGCTGTCCCTCCGCTCTCTAT	3217
Db	3358	GGGAAGCAAAACCGCACCGGACCAATTTGCTGTTGCGGTATCTCTCTTGGCTATCTCTCTGAT	3417
QY	3218	CTTCTCTCTCTGCTGGGTTCTGTTGATCCATTCACTACCGTCTCGCTGGCCCAAAATAT	3277
Db	3418	CTTTTCTCTGATGGGTTCTGATCGATCCATTCACCAACCGGGTCACTGGCCCTGATAT	3477
QY	3278	CCAAACCTGTGGCATCAACTGCTAGGA	3304
Db	3478	CGCGAAATGTGGCATCAACTGCTAGGA	3504
RESULT 6			
AAZ58271			
ID	AAZ58271 standard; cDNA; 1734 BP.		
AC	AAZ58271;		
DT	08-MAY-2000 (first entry)		
DE	Wheat cellulose synthase cDNA contig.		
XX	Wheat; cellulose synthase; transgenic plant; ss.		
OS	Triticum aestivum.		
FT	Key	Location/Qualifiers	
FT	CDS	2..1522	
FT		/*tag= a	

FT XX /partial
PN XX WO200004166-A2.
XX XX 27-JAN-2000.
PD XX
PF XX 13-JUL-1999; 99WO-US15871.
PR XX 14-JUL-1998; 98US-0092844.
XX XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX XX
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;
PI Rafalski JA, Thorpe CJ;
XX XX
DR WPI: 2000-182431/16.
DR P-PSDB: AAY58840.
XX XX
XX Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as
PT probes for isolating cDNAs and genes encoding homologous proteins, for
PT producing transgenic plants -
PS Claim 3: Page 66-67; 93pp: English.
XX XX
CC The present sequence is that of a contig of cDNA clones and encodes
CC a portion (see AAY58840) of wheat cellulose synthase (CS). The cDNA
CC clones were isolated from wheat 7-day seedling leaf and 7-day
CC etiolated seedling root tissue cDNA libraries on the basis of
CC homology to Arabidopsis and cotton CS sequences. The invention
CC relates to isolated nucleic acid fragments encoding plant CS and to
CC CS polypeptides. It also relates to the construction of a chimeric
CC gene encoding all or a portion of the CS, in sense or antisense
CC orientation, where expression of the gene results in altered levels
CC of the CS in transformed host cells. The host cells can be used to
CC screen compounds for their ability to inhibit CS activity. CS
CC nucleic acids are also useful for producing transgenic plants having
CC altered levels of CS, and hence altered levels of fibre. CS may
CC also serve as a target for the development of novel herbicides.
XX XX
XX Sequence 1734 BP; 441 A; 369 C; 420 G; 502 T; 2 other;
Query Match 46.4%; Score 1681.6; DB 21; Length 1734;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1782 GAAGCTATGTCTCCCTAATGGATCCAAACCTAGTCCGCAAGTCTGTATGTGCAGTTC 1841
Db 2 GAAGCTATGTCTCCCTAATGGATCCAAACCTAGTCCGCAAGTCTGTATGTGCAGTTC 61
QY 1842 CCACAAAGGTTTGATGGATTGATAGGAATGATCGATATGCAACAGAACACTGCTTT 1901
Db 62 CCACAAAGGTTTGATGGATTGATAGGAATGATCGATATGCAACAGAACACTGCTTT 121
QY 1902 TTTGATATTAACCTTGAAGGCGCTTACGGGATTCAGGACCACTTTATGTGGAACTGCT 1961
Db 122 TTTGATATTAACCTTGAAGGCGCTTACGGGATTCAGGACCACTTTATGTGGAACTGCT 181
QY 1962 TGTGTTTTCACAGACAGATCTATCTATGTTATGAGCCCAATTAAGCGGAAGACCA 2021
Db 182 TGTGTTTTCACAGACAGATCTATCTATGTTATGAGCCCAATTAAGCGGAAGACCA 241
QY 2022 GGTTCCTTGGCATCCTATGTGGGGCAAGAAGCAAGCAAGTCAAGAAAGAGGAGC 2081
Db 242 GGTTCCTTGGCATCCTATGTTGGGGCAAGAAGCAAGCAAGTCAAGAAAGAGGAGC 301
QY 2082 TCAGATAGAAAGTGCACACAGCATGTGGACAGTTCTGCCAGTTTCAATCTCGAA 2141
Db 302 TCAGATAGAAAGTGCACACAGCATGTGGACAGTTCTGCCAGTTTCAATCTCGAA 361
QY 2142 GACATAGAGGAGGTTGTTGAAGTGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCT 2201
Db 362 GACATAGAGGAGGTTGTTGAAGTGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCT 421

QY 2202 CAATGAGCTTTAGAGAAGAGAT 2261
Db 422 CAATGAGCTTTAGAGAAGAGAT 481
QY 2262 GAATATGTTGTTGTTCTCTCAGT 2321
Db 482 GAATATGTTGTTGTTCTCAGT 541
QY 2322 GTCATAAGTTGTGGCTATGAGG 2381
Db 542 GTCATAAGTTGTGGCTATGAGG 601
QY 2382 GGATCTGTACAGAGATATTC 2441
Db 602 GGATCTGTACAGAGATATTC 661
QY 2442 GTCATTATGTCATGCCCAAGCGCC 2501
Db 662 ATCTATTGTCATGCCCAAGCGCC 721
QY 2502 CGTCTGAACCAAGTGTGCGGT 2561
Db 722 CGTCTGAATCAAGTGTGCGGT 781
QY 2562 TGCCCTTATGTTATGGCTACG 2621
Db 782 TGCCCTTATGTTATGGCTACG 841
QY 2622 AACACACCAATTTACCACCTAA 2681
Db 842 AACACACCAATTTACCACCTAA 901
QY 2682 ATCTGTCTGCTCACTGGAAAGT 2741
Db 902 ATCTGTCTGCTCACTGGAAAGT 961
QY 2742 TTCATTGGGCTTTCTTTCAA 2801
Db 962 TTCATTGGGCTTTCTTTCAA 1021
QY 2802 GTTGGCATTCAGAGTGTGGAG 2861
Db 1022 GTTGGCATTCAGAGTGTGGAG 1081
QY 2862 CATCTGTTTGGCGCTTTTCAGG 2921
Db 1082 CATCTGTTTGGCGCTTTTCAGG 1141
QY 2922 ACTGTCACCTCAAGGCTAATG 2981
Db 1142 ACTGTCACCTCAAGGCTAATG 1201
QY 2982 TGGACGAGCTTCTCATCCCTCG 3041
Db 1202 TGGACGAGCTTCTCATCCCTCG 1261
QY 3042 GCTGGACCTCTACGCCATCA 3101
Db 1262 GCTGGTACCTCTACGCCATCA 1321
QY 3102 CTCCTTCTTGGCTTCTGGGTGAT 3161
Db 1322 CTCCTTCTTGGCTTCTGGGTGAT 1381
QY 3162 AGGCAAAACCGCACACCGAGAT 3221
Db 1382 AGGCAAAACCGCACACCGAGAT 1441
QY 3222 TCCTTGTCTGGGTTCTGTTTGA 3281
Db 1442 TCCTTGTCTGGGTTCTGTTTGA 1501
QY 3282 ACCTGTGGCATCAACTGCTAGG 3341

```

Db 1502 ACCGTGGCATCACTGCTAGGAAGTGGAGTTTGTAGACACAGAAATATAACAGTGA 1561
QY 3342 TCGAGCAACAACCCGGGAGCAGAGATATTTATGTTGGGTTGTGAATTACTACGTTT 3401
Db 1562 TCGAGCGCACCTGTGGAGCCAGAGATATTTATGTTGGGTTGTGAATTACTACGTTT 1621
QY 3402 GAGAAGTTGTCAAATTCAGAAAAACACATTTGTAATAGATGTAATAGACCATCTACCG 3461
Db 1622 GAGAAGTTGTCAAATTCAGAAAAACACATTTGTAATAGATGTAATAGACTAICTACCG 1681
QY 3462 TTTTCATGAGGTTAAGCTCTTCTTTTGGAAACAAAGGAA 3502
Db 1682 TTTTCATGAGGTTAAGCTCTTCTTTTGGAAACAAAGGAA 1722

RESULT 7
AAC49550
ID AAC49550 standard; DNA; 3198 BP.
XX
AC AAC49550;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
XX
PF 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
```

PR	05-AUG-1999;	99US-0147192.	Query Match	46.39;	
PR	05-AUG-1999;	99US-0147260.	Best Local Similarity	72.08;	
PR	06-AUG-1999;	99US-0147303.	Matches 2330; Conservative		
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	23-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
PR	07-SEP-1999;	99US-0152363.			
PR	10-SEP-1999;	99US-0153070.			
PR	13-SEP-1999;	99US-0153758.			
PR	12-SEP-1999;	99US-0154018.			
PR	16-SEP-1999;	99US-0154039.			
PR	20-SEP-1999;	99US-0154779.			
PR	22-SEP-1999;	99US-0155139.			
PR	23-SEP-1999;	99US-0155486.			
PR	24-SEP-1999;	99US-0155659.			
PR	28-SEP-1999;	99US-0156458.			
PR	29-SEP-1999;	99US-0156596.			
PR	04-OCT-1999;	99US-0157117.			
PR	05-OCT-1999;	99US-0157753.			
PR	06-OCT-1999;	99US-0157865.			
PR	07-OCT-1999;	99US-0158029.			
PR	08-OCT-1999;	99US-0158232.			
PR	12-OCT-1999;	99US-0158369.			
PR	13-OCT-1999;	99US-0159293.			
PR	13-OCT-1999;	99US-0159295.			
PR	14-OCT-1999;	99US-0159329.			
PR	14-OCT-1999;	99US-0159330.			
PR	14-OCT-1999;	99US-0159331.			
PR	14-OCT-1999;	99US-0159637.			
PR	14-OCT-1999;	99US-0159638.			
PR	18-OCT-1999;	99US-0159584.			
PR	21-OCT-1999;	99US-0160741.			
PR	21-OCT-1999;	99US-0160767.			
PR	21-OCT-1999;	99US-0160768.			
PR	21-OCT-1999;	99US-0160770.			
PR	21-OCT-1999;	99US-0160814.			
PR	21-OCT-1999;	99US-0160815.			
PR	22-OCT-1999;	99US-0160980.			
PR	22-OCT-1999;	99US-0160981.			
PR	22-OCT-1999;	99US-0160989.			
PR	25-OCT-1999;	99US-0161404.			
PR	25-OCT-1999;	99US-0161405.			
PR	25-OCT-1999;	99US-0161406.			
PR	26-OCT-1999;	99US-0161359.			
PR	26-OCT-1999;	99US-0161360.			
PR	26-OCT-1999;	99US-0161361.			
PR	28-OCT-1999;	99US-0161920.			
PR	28-OCT-1999;	99US-0161992.			
PR	28-OCT-1999;	99US-0161993.			
PR	29-OCT-1999;	99US-0162142.			

Qy	77	CGCCCTGAAGTCCGGAGGCAC	are 1677.6; DB 21; Length 3198;
Db	21	CGCGGAAAGCCGATGAAGAAC	ed. No. 0;
Qy	137	GGGCACACGTTGGACGGGAC	Mismatches 834; Indels 72; Gaps 8;
Db	81	TGGCAAGACTGTTGATGGAGAT	
Qy	197	CGGCCCTGCTACGAGCAGCAG	CGGGGACGTGTGCAGATCTCGCCGACGGCCT 136
Db	141	TCGGCCTTGCTACGAGTATGAG	
Qy	257	CAAGTACAGGCCACAGAGGG	TCGCAGACTTGCAGATCTGTAGTACAATGT 80
Db	201	CAGATACAGAGGCTCAAGGT	CACCGCTGGCAGCTGCGCGCTTCCCGGCTG 196
Qy	317	TGCCGATGATGTA---GTGAC	
Db	261	AGCTGATGAAGTACTGTTGAG	GTGCGCTTGATATTTGTTTCATTTCCAGTTG 140
Qy	374	GAAGATTGCTACAGGATGCGC	GGAGGCAACCAGGCCCTCGCTCCAGTCAAGAC 256
Db	303	GAATTTTCAGAGCGGATGCTT	AGATGGGAATCAATCTTGTCTCAGTGCACAAAC 200
Qy	434	CCACCCCAAGTATCACAGTGGC	AGGATCCGGGGGAGGAAAGCGGACGACACTGA 316
Db	363	GGAACCCAGTATGATAAAGAG	
Qy	494	TAGGGATACGTCCTTCAGTC	GTCTATTCTCGGTGATTAAGACGAGGATGGCTT 260
Db	414	GAGCAGACAAG-----	CTACCCCTGCATCTGGCACTGAGGACCAAGCA 373
Qy	554	GCCTGATCATCATGATGTCC	CTACCCCT-----CAGAAGCA 302
Db	453	ACCTGAACGCTCTCTGTATCT	GGCATGAACCCGGGGGCGAGTGGCAATGTTGG 433
Qy	613	--ATGTGAATCATTCACCAAT	GCATCTTACTCGTGGGAAAGGAGAGAAATGGG 362
Db	513	AGATGTCAATCAATCACCAAAT	CGGCCCTCCAAGTATGACAGTGGAGAGATCCC 493
Qy	671	CTGGAAGAGAGAGTTGATGGCT	-----TCTCTCAATCATCTTCTCGTCTCAC 413
Db	573	TTGGAAGGAGAGCTTGATGGCT	CAGCCAGATGTCAGGAGAAATCCTTGGAGCTTC 553
Qy	731	TAATGGGACAGCAATTGCTCCCT	-----ATACTTCAGGAGAGTTTCTGCTGCTC 452
Db	633	CACGCAG-----GCTGCT	GGGGAACATCAGCAGAGCTGCTCCGTTTCCT - 612
Qy	791	TGAATACAACATGGAAGACGCT	TATCGCTGGGGAAAGCGCTTCCCTATTATCATC 512
Db	681	AGATATCCTAGCAGATGAGGCT	GAAGGAGTCTCCGGCAGTATTGGGAATGTTGC 670
Qy	851	AGTCCCATTTGCTTCTTCCCAAA	GATTGGATCTCTGGACTCGGGAATGTAGC 572
Db	741	AGTTTCAATTCCTTCATCACGGA	AATGAAGCAGCAAGAGGTGCGATTCCCATGAC 730
Qy	911	TGTTCTAAGCATCTTCTTGCAC	AATGAAGCAAGAGAAATACTGCTCTGTGAC 632
Db	801	TATCCTTTGTCTCTTCTTGCAT	AGTCTGGGAGCTACTGACATCGATGATCTAC 790
Qy	971	GTGCTTTTATCTCTTATATGTG	AAGAGTGGAG---TAGATATTGATGCCAGCAC 680
Db	861	ATGGCTGGTCTCTGTATATGTG	GAATGATGAACCTGCCAGCCTCTATCTAGAAA 850
Qy	1031	CCCGAAGTGGTTTCCCAATCAAC	GAATGAGAAAGCGGAGGAGCCTCTGTCAAGGAA 740
Db	921	TCCCAAGTGGTTTCTCTGTGAAC	TCCTTACAGAATGGTCAATGTTTGGGGTGGT 910
Qy	1091	CCGAGAAGTGAACCGTCTCAGT	TCCTTACAGAATGGTTATTATGCTGCGGCTGT 800

TCGTGTTGACATATTTGTCAGTACAGTCGACCC 1150

Db	981	TCGTGAAGGTGAGCCATCACAGTTAGCTGCTGTGGACATTTTCGTGAGTACTGTTGACCC	1040
Qy	1151	CTTTGAAGGAGCCACCTATTCGCTCAGTCGCCAACACTGTGCTATCCCATTCCTGCTGTTGATTA	1210
Db	1041	CTTTGAAGGAGCCACCCCTGTGTGACAGCCCAACACAGTGTCTCTATTCTGGCTGTGACTA	1100
Qy	1211	TCCCGTGGACAGGTCCTTTGGTATGTATCTGATGACGGAGCTTCAATGCTGACTTTTGA	1270
Db	1101	CCCAGTTGACAGGTCCTCGTTATGTTTCTGTATGATGGTCTGCTATGTTATCAUATTGA	1160
Qy	1271	CGCATTTGGCTGAGACTTCAGAGCTTTCGCTAGGAATGGGTACCATTTGTGGAAGAGTATGA	1330
Db	1161	ATCAGTTGCAGAAACATCAGAGTTTGTCTGTAATGGGTACCATTTTGCAGAAATATAG	1220
Qy	1331	CATTGAACCCAGAGCTCCCGAGCTTTTACTTTTGGCAGAAAAATGATTACCTGAAAGACAA	1390
Db	1221	CATAGAGCCTCGTGCACCAAGATGTTACTTTGCTGCGAAAAATAGATTACTTTGAAGGATAA	1280
Qy	1391	AGTCCAGCCTTCATTTTGTAAAGACCGCGGCCCATGAAGAGAGAAATATGAAGAAATTTAA	1450
Db	1281	AGTTCAGACATCATTTGTCAAAAGATCGTAGAGCTATGAAGAGGAAATATGAGGAATTTAA	1340
Qy	1451	AATCAGGATAAATGCCCTAGTTTCTTAAGGCATTGAAGCTCCCGCAGGAAGATGGATCAT	1510
Db	1341	AATCCGAATCAATGCACCTTTGTTCCAAAGCCCTAAATGTCTCTGAAGAGGTTGGTTAT	1400
Qy	1511	GCAAGATGGCACACATGCGCAGGAAACAATACCAGGATCATCTCGAATGATTCAGGT	1570
Db	1401	GCAAGATGGCACACCGTGGCCTTGGAAATTAATCAAGGACCATCCAGGAATGATCCAGGT	1460
Qy	1571	TTTCCCTTGGTCACAGTGTGGCTTTGTATCTAGGGTAATCAGCTCCCCGCTTTAGTTTA	1630
Db	1461	CTTCTTAGGCCAAATGGTGGACTGTATGCAGAGGSCAATGAGCTCCCGCTTGGGTATA	1520
Qy	1631	TGTGTCCTGTGAAAGGCTCTGGGTTCAGCACACAAAGAGCTGGTGCCATGAATGC	1690
Db	1521	TGTTTCTCGAGAAAGCGACCAAGATTCAGCACCAAAAAGGCTGGTGTATGAATGC	1580
Qy	1691	CCTTGTTCGTGCTCAGCTGTCCTTTACTAATGGAACAATACATGTTGAATCTTGATCTGGA	1750
Db	1581	ACTGTTGAGAGTTTCAGCAGTTCTTACCNAITGGACCTTTCATCTTGAATCTTGATGTGA	1640
Qy	1751	TCACTACATCAACAACAGCAGGCTGTCCGAGAAGCTATGTCTTCCTCAATGATGCCAAA	1810
Db	1641	TCATTACATAAATAACAGCAAAAGCCCTTAAGAGAAGCAATGTCTCTGTATGGACCCAAA	1700
Qy	1811	CCTAGTCCGCAAGTCTGTTATGTCCAGTTCCCACAAAGGTTTCATGGGATTCATAGGAA	1870
Db	1701	CCTCGGAGCAGTGTGTTATGTTCAGTTCCCACAAAGATTTGATGGTATPCGATAAGAA	1760
Qy	1871	TGATCGATATGCAACAGGAACACTGTCTTTTTCATATTAACITGAGGGCCTTGACGG	1930
Db	1761	CGATAGATATGCTAATCGTAATACCGTGTCTTTCATATTAACITGAGAGGTTTAGATGG	1820
Qy	1931	CATTCAAGNACCAGTTTATGTGCGGAAGCTGGTTGTGTTTTCAACAGAACACCTATCTATGG	1990
Db	1821	GATTCAAGGACCTGTATATGTCCGAAGCTGGATGTGTTTTCAACAGAACACGATTAACGG	1880
Qy	1991	TTATGAGCCCCCAATTAAGS-----CGAAGAAGCCAGGTTTCTTGGCATCACTATGTGG	2044
Db	1881	TTATGAAGCTCCAATAAAGTAACACAGAAGCCAGTCTTTTATCTAAGCTCTGTGG	1940
Qy	2045	GGCAGAGAGGAGGCAAGCAATCAAGAAAGGAGCTCAGATAGAAAAAGTCGCAACAA	2104
Db	1941	TGATCAAGAAAGAAAGAAATTCCAAAGCTAAGAAAGAGTCGACAAAAAGAAATCAGCAG	2000
Qy	2105	GCATGTGGACAGTCTGTGTTCCAGTATTCATCTCGAACACATAGAGAGGCTGTGTAAGG	2164
Db	2001	GCATCTGACTCAACTGTTCTCTGTATTCACCTCGATGCATAGAAGAGGAGTTTGAAGG	2060
Qy	2165	TGCTGGSTTTGATGATGAGAAATCAGTTTCTCATGTGCTCAAAATGAGCTTAGAGAGAGATT	2224

Db	2061	TGCTGGTTTTGATGATGAAAGGGCGCTCTTAATGTCGCAAAATGAGCCTGGAGAAAGCGATT	2121
Qy	2225	TGGCCAGTCACGACGACATTTGTTGCCCTCCACTCTGATGGAATATGTTGTTCTCTCASTC	2284
Db	2121	TGGACAGTCGCTGTTTTGTTGCTTCTACCCCTAATGGAATAATGTTGTTCTCTCTTC	2180
Qy	2285	CTCCACTCCAGATCTCTTTTGAAGAAGCTATCCATGTCTATAGTTTGTGCTATGAGGA	2344
Db	2181	AGCAACTCCAGAAACACTTCTCCAAAGAGGCTATCCATGTCTATAGTTGTTGTTATGAGGA	2240
Qy	2345	CAAGTCTGAATGGGGAACTGAGATTGGTTGGATCTATGGATCTGTACAGAAAGATATTCT	2404
Db	2241	TAACTCAGATTGGGAATGGAGATTGGATGGATCTATGGTTCTGTGACAGAAGATATTCT	2300
Qy	2405	TACTGGATTCAAGATGACGACAGAGGCTGGGCTTCACTTATTGCTATGCCCCAAGCGCCC	2464
Db	2301	GACTGGGTTCAAAATGATGCCCGTGGATGGCGGATCCATTTACTGCTATGCTTAAGCTTCC	2360
Qy	2465	AGCTTTCAAGGGATCTGCCCCATCAATCTTTTCAGATCGTCTGACCAAGTCGTCGGTG	2524
Db	2361	AGCTTTCAAGGGTTCTGCTCTATCAATCTTTTCAGATCGTCTGAAACCAAGTCGTGAGGTG	2420
Qy	2525	GGCTCTCGGTTCTGTTGAAATCTTTTCAGCCGGATATGCCCTTATGGTATGGCTACGG	2584
Db	2421	GGCTTAAAGTTCAAGTTCCTCTTCAGTCGCGCATGCTCTATATGGTATGGTTACAA	2480
Qy	2585	AGGGCGCTCAAGTTCCTGGAGAGATTCCGTTACATCAACACCACCATTTFACCACATAC	2644
Db	2481	TGGGAGGCTAAAATTTCTTGAGAGGTTTCGGTATGTGAACACCACCATTACCCTATCAC	2540
Qy	2645	CTCTCTCCGCTTCTAGTCTATTGTATATTGGCTGCTATCTCTGCTCACTCGAAAAGTT	2704
Db	2541	CTCCATTCCCTCTCATGTATTGTACATTGCCAGCCGTTTGTCTCTTCACCAACCAGTT	2600
Qy	2705	CATCATGCCAGAGATTAGCAACTTGGCCAGTATCTTGGTTTCATTGGGCTCTTCTTTCAT	2764
Db	2601	TATTATTCCCTCAGATTAGTAACATTCGAAGTATATGGTTCTGTCTCTCTTCTCTCCAT	2660
Qy	2765	TTTCGCACCTGGTATCCCTTCAGATGAGTTGAGTGGTGTGGCATGACAGTGGTGGAG	2824
Db	2661	TTTCGCCACGGGTATCTAGAAATGAGTTGGAGTGGCGTGGCATAGACGAATGGTGGAG	2720
Qy	2825	GAATGAACAGTCTGGGTCAATGGAGGTATCTCTGCACATCTGTTTGGCGCTTTTCAGGG	2884
Db	2721	AAAGGAGCAGTTTTGGGTCAATTCGTTGGAGTATCCCGCTCATTTATTTCGTGTGTTTCAAGG	2780
Qy	2885	TCTTCTGAAGTGCTTCGCGGTATCACACCAACTTCAGCTGTCACCTCAAAAGCTAATGA	2944
Db	2781	TATCCTCAAGTCTCTGCCGGTATTGCACAAACTTCACAGTTACCTCAAAAGCTTCAGA	2840
Qy	2945	CGAAGAAGGCGACTTTTGCTGAGCTCTACATGTTTCAAGTGGAGCGAGCTTCTCATCCCTCC	3004
Db	2841	TGAAGCGGAGACTTTGCTGAGCTCTACTTGTTCAAATGGACAACACTTCTGATTCCGCC	2900
Qy	3005	GACGACCATTTTGATCAATTAACATGTTGGTGTGCTGCTGGCAGCTCCTTACGCCATCAA	3064
Db	2901	AACGACGCTGCTCAATGTAACCTTAGTGGGAGTGTGTGCAGAGTCTCTTAGTCTATCAA	2960
Qy	3065	CAGTGGTTTACCAATCATGGGGGCGCTCTTTTGGGAAGCTCTTCTTGTGCTTTCTGGGTGAT	3124
Db	2961	CAGTGGATACCAATCATGGGGACCACTTTTGGTAAGTTGTTCTTGTGCTTCTGGGTGAT	3020
Qy	3125	TGTTCACTTATACCAATTCCTCAAGGGTCTTATTGGCAGGGCAAAACCCACACCGACGAT	3184
Db	3021	TGTTCACTTGTACCTTTCTCAAGGGTTTGATGGGTGCAGAGAACCGGACTCTTACCAT	3080
Qy	3185	TGTCATCTCTGCGCTCTCTCTCTGCTCTATCTTCTCTCTGCTGTTGGGTTCTGTGTG	3244
Db	3081	TGTTGTGGTCTGGTCTGTTCTCTCTCTGCTCTATCTTCTGTTGTGGGTTAGGATTGA	3140
Qy	3245	TCCATTCACTACCGGCTCTGCTGGCCCCAAATATCCAAAACCTGTGGCATCAACTGCT	3300
Db	3141	TCCCTTCACTAGCCGAGTCACTGGCCGGACATCTTGGAAATGGAACTCAACTGTT	3196

QY 1250 AGCTTCAATGCTGACATTTTGACGCATTTGGCTGAGACTTCAGAGTTTGTGCTAGGAATGGGT 1309
D 1351 CTGAGCTATGCTGACATTTTGTAGTCTCTCTGAAATGCTCGCAATTTGTAGAAAGTGGGT 1410
QY 1310 ACCATTTTGAAGAAGATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTGGCCAGAA 1369
D 1411 TCCTTTTGAAGAAGCAATATTTGAACCAAGAGCTCCAGAAATTTACTTTTGTCTCAAAA 1470
QY 1370 AATTGATTACCTGAAAGACAAAGTCCAGCCTTCATTTGTTAAAGACCCCGCCGCATGAA 1429
D 1471 AATAGATTACCTGAAGGACAAAATCAACCTTTCATTTGTTAAAGGAAAGACGAGCAATGAA 1530
QY 1430 GAGAGAAATGAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTTAAAGT 1489
D 1531 GAGAGAGATTGAAGAATTCARAATAAGAAATCAATGCCCTTGTGGCAAGACACAGAAAGT 1590
QY 1490 CCCGAGGAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAACAAATPACGAGGA 1549
D 1591 GCCTGAAGAGGGGTGGACCATGCTGATGGAATGCTTGGCCTGGGAATAACCCTAGGGA 1650
QY 1550 TCATCTGGAAATGATTCAGGTTTTCCTGGTCACAGTGGTGGCCTTGATACGTAGGGTAA 1609
D 1651 CCATCTGGCATGATTCAGGTTGTTCTGGGGCACAGTGGTGGCCTTGACACTGTATGGAAA 1710
QY 1610 TGAGCTCCCGCTTAGTTTATGTGTCCTGTAAGAGCCTCCTGGGTTCCAGCACACAAA 1669
D 1711 TGAATACCAAGTCTGTCTATGTCTCTGCTGTAAGAGACACAGGCTTTCAGCATCACAA 1770
QY 1670 GAAGGCTGGTGCATGAATGCCCTTGTCTGCTCAGCTGCTCCTTACTTAATGGCAATA 1729
D 1771 GAAGGCTGGTGCATGAATGCCACTGATTCGTGTATCTGCTGCTGACAAATGGTGCCTA 1830
QY 1730 CATGTTCAATCTTGAATGTGATCAGTACATCAACACACAAAGGCTGCCGAGAGCAAT 1789
D 1831 TCCTTCAAGTGGATGTGACCATTACTTCAATAGCAGCAAGCTTTAGAGAGCAAT 1890
QY 1790 GTGCTTCTTAATGGATCCAAACCTAGGTCCGCAAGTCTGTATGTGACGTTCCCAACAAAG 1849
D 1891 GTGCTCATGATGCATCCAGCTCTAGGAAGGAAAATCTGTATGATACAAATTTCCACAAAG 1950
QY 1950 GTTTGATGGATGTAGGAAATGATCGATATGCAACACAGACACTGCTTTTGTGATAT 1909
D 1951 ATTTGATGGCAATTTGACTTTGCACGATCGATGCTAATAGGAACATAGTCTTCTTTGATAT 2010
QY 1910 TAATTTGAGGGCCTTGACGGCATTCAGGACCAATTTATGTGGGAACGTTGTGTTT 1969
D 2011 CACATGAAAGGCTTAGATGGCAATTCAGGTCGAGTCTATGTGGGAACAGGATGCTGTT 2070
QY 1970 CAACAGAACAGCTATCTATGGTTATGAGCC---CCCAATTAAGCGGAAGACCGAGTTT 2026
D 2071 CAATAGCAGGCTTTGTATGGATATGATCCTGTTTGTGACTGAAGCTGATCTGGAACCTAA 2130
QY 2027 CTGCGCATCACTATGTGGGGGCAAGAAAGGCAAGCAAGTCAAGAAAAGAGCTCAGA 2086
D 2131 CATTTGTTTGAAGAGCTGCTGTGGTGAAGGAAGAGAAAGCAAGAGTTATATGGATAG 2190
QY 2087 TAAGAAAAAGTCGAACAGCATGTGGACACTTCTGTTCCAGTATTTCAATCTCGAAGACAT 2146
D 2191 TCAAGCCGTTATTTGAAGAGACAGAAATCTTCAGTCCCATCTTTAAGATGGAAGACAT 2250
QY 2147 AGAGGAGGGTGTTCAGAGTCTGGGTTTGATGATGAGAATCAGTTCTCATGTCTCAAAAT 2206
D 2251 CGAGGAGGGTATTGAAGGT-----TATGAGGATGAAAGTCAAGTCTTATGTCCAGAG 2304
QY 2207 GAGCTTAGAAGAGATTTGGCCAGTCAGCAGCATTTGTGGCTCCACTCTGATGGAATA 2266
D 2305 GAAATTTGGAGAAAGCCTTTGGTCAGTCTCCAAATCTTCATTTGCATCCACCTTTATGACTCA 2364
QY 2267 TGTGTGTCTTCCCTCAGTCTCCATCCAGAACTCTCTTTTGAAGAGAGCTATCCATCTCAT 2326
D 2365 AGGTGGCATCCACCTTCAACAAACCCAGGCTTCTTACTGAAGGAAGCTATCCATGTTAT 2424
QY 2327 AAGTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGGTTGGATCTATGGATC 2386

D 2425 CAGCTGTGGGTAGAGGACAAAATGAATGGGAAAAGAGATTTGGCTGGATCTATGGTTC 2484
QY 2387 TGTACAGAAAGATATTTACTTACTGATTTCAAGATTCAGCGCAGAGGCTGGGCTTCACTCA 2446
D 2485 AGTTACAGAGATATTTCTGACTGGTTTAAATGATGCAAGAGGCTGGCAATCAATCTA 2544
QY 2447 TTGCATGCCCAAGCGGCCAGCTTTCAAGGATCTGCCCCATCAATCTTTTCAGATGGTCT 2506
D 2545 CTGCATGCCCAAGCAGACTTTGTTTCAAGGGTTCTGCACCAATCAATCTTTCTGATGGTCT 2604
QY 2507 GAACCAAGTGTGCGGTGGCTCTCGGTTCTGTGTGAAATTTCTTTTTCAGCCGGCATTTGCC 2566
D 2605 TAATCAGGTGCTCGTTGGGCTCTTGGTCAAGTGGAAATTTCTGCTAGCAGACATTTGTC 2664
QY 2567 CTATGATGATGGCTACGAGGAGGCGCTCAAGTTCTCTGGAGAGATTCGCTTTACATCAACAC 2626
D 2665 TATATGATGGCTACATTTACATTTTATCAATTTTATCTGAGATTTAGTAAATATGCTGGAATTTCTTCAT 2724
QY 2627 CACCATTTACCCACTAACCTCTCTCCGCTCTAGTCTATTGTATATTGCTGCTATCTG 2686
D 2725 CATTTGTTATCCAAATCAATCTGTTCCGCTTATCGCCTTATGCTGCTTCTGCTATCTG 2784
QY 2687 TCTGCTCAGCTGGAAGTTTCATTCATGCGCAGAGATTTAGCAACTTTGCCAGTATCTGTTTCAT 2746
D 2785 TCTTCTTACCAATAAATTTATCAATTTTATCTGAGATTTAGTAAATATGCTGGAATTTCTTCAT 2844
QY 2747 TGCCTCTTCTTCAATTTTTCGCACTGGTATCTTTCAGATGAGGTGGAGTGGTGG 2806
D 2845 TCTTCTTTTTCGCTCCATTTTTCGCACTGGTATATTTGAGAGCTCAGATGGAGTGGTGG 2904
QY 2807 CATTTGACAGTGGTGGAGGAATGAACATTTCTGGGTCATTTGGAGGTATCTCTGCACATCT 2866
D 2905 CATTTGAAGATTTGTTGGAGAAATGAGCAGTTTGGGTTATTTGGTGGCACCTCTGCCATCT 2964
QY 2867 GTTTGCCGCTTTTTCAGGTCCTTCTGAAAGTGTTCGCGGTATTCGACACCACTTCATCTGT 2926
D 2965 CTTCGCGGTGTTCCAGGCTCTGCTGAAAGTGTTCGCTGGGATTTGATACCACTTCACAGT 3024
QY 2927 CACCTCAAGAGCTTAATGACGAAGAAGCGACTTTGCTGAGCTCTACATGTTTCAAGTGGAC 2986
D 3025 TACCTCAAGAGCATCTGATGAGGATGGCAGCTTTTGTGAGCTATATGTGTTCAAGTGGAC 3084
QY 2987 GACGCTTCTCATCCTCCGAGCAGCATTTTTCATTAACATGTTGTTGCTGTTGCTGG 3046
D 3085 CAGTTTGTCTATCCTCCGACCACTGTTCTTGTCTATTAACCTGCTCGGAATTTGGSCAGG 3144
QY 3047 CACTCTCTACGCCATCAACAGTGGTTACCAATCATGCGGGCGCTCTTTTGGGAAGCTCTT 3106
D 3145 AATTTGCTATGCCATTAACAGCGGCTACCAATCCTCGGGGTCCGCTCTTTGGAAGCTGTT 3204
QY 3107 CTTTTCCTTCTGGGTGATTTGTTTACTTATACCATTTCTCAAGGCTCTTATGGGCAAGCA 3166
D 3205 CTTTTCGATCTGGGTGATCTCCATCTCTACCCCTTCTCAAGGCTCTCATGGGCAAGCA 3264
QY 3167 AAACGCGCACCGCAGGATTTGCTGCTGGGCTGCTCCTCCTGCTTCTATCTTCTCCTT 3226
D 3265 GAACGCGCAGCAACAAATCGTCTGTTGTTGCTCCATCTCCTTGGGCTATCTTCTCCTT 3324
QY 3227 GCTGTGGGTTTCGTTTGTATTCATTCACATACCC 3258
D 3325 GCTGTGGGTGAAGATCGATCTCTTTCATCTCCC 3356

RESULT 10
AAZ99506
ID AAZ99506 standard; DNA; 3725 BP.
XX
AC AAZ99506;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.

XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX
OS Zea mays.
FH Key Location/Qualifiers
FT CDS 179..3397
FT /*tag= a
FT /product= "cellulose synthase"
FT /note= "no termination codon given"
XX
PN WO200009706-A2.
XX
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US18760.
XX
XX 17-AUG-1998; 98US-0096822.
XX
XX (PION*) PIONEER HI-BRED INT INC.
XX
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
PI P-PSDB; MAY84112.
DR WPI; 2000-224343/19.
DR P-PSDB; MAY84112.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein -
XX
PS Claim 1; Page 121-126; 119pp; English.
XX
XX The present sequence encodes a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;

Query Match 38.88; Score 1408.4; DB 21; Length 3725;
Best Local Similarity 66.88; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;

QY 53 CGCGCCCATGGACGGCGGACGGCCCTGAAGTCCGGGAGGCACGGGCGGGGACGT 112
DB
DB 238 CGGCCACGACGGCGGCGGCTGTCCCGGCTAAGCCACGAAGAGTGCGAATGGCGAGT 297
QY 113 GTGCCAGATCTGCGCGGACGGCCCTGGGACACCTTGGAGGGGACGTCTTACCGCGCTG 172
DB 298 GTGCCAGATTTGTGGCGACACTGTGTGGCGTTTCAGCCCACTGGTGATGCTTTTGGCTG 357
QY 173 CGACCTCTCGCGCTTCCCGGCTCTCCGCCCTGTCTACGACACGAGCGCAAGGAGGCAC 232
DB 358 CAATGAGTGTGCCCTTCCCTGTCTCGCCGCCCTTGTATGATGATGAGCGCGCAAGAGGAA 417

QY 233 CCAGCGCTTCCCTCCAGTGCAC 292
DB 418 CCAATGCTCTGCCCTCAGTGCAC 477
QY 293 CGGGGAGGAAGGCGACGAC 352
DB 478 TGGTATGATGAGGAGAGAT 537
QY 353 TGGCACTGAGGACAGAGAC 412
DB 538 CAATGGGAGGCGCCAGAGTG 574
QY 413 CGGGGCGAGTGCATGTTGG 472
DB 575 GACGTGATCTCTTCTCATCTG 634
QY 473 GTATGACAGTGGAGAGATCCC 532
DB 635 AGTGGACA-----ACAGATATC 651
QY 533 AGGAGAAATCCCTGGAGCTTCG 592
DB 652 TGGAGAGATCCCTGATGATCG 711
QY 593 CAGACGTGCTCGTTCCTCCAT 649
DB 712 TGTTCATCAAGCGTTCAGTTC 771
QY 650 CGGCAGTATGGGAATGTTGGC 709
DB 772 CTATGGGCTTAATAGTGTGAC 831
QY 710 CAAGGGTTCGATTCCTCCATGAC 769
DB 832 CAAAAATATGTTGCAAGTGACT 870
QY 770 TACTGACATCATGATCATCTACT 829
DB 871 TAGAGAGACATGGAGGGGACT 930
QY 830 TCGCCAGCGCTCTATCTAGAAAA 889
DB 931 AGCCCTACCTTTGAGCCCAAT 990
QY 890 GGTGATGTTCTGCGGTTGGTTC 949
DB 991 AGTAATCATCTCCGCTCTTATC 1050
QY 950 TGTGCGTAATCATACCCACTG 1009
DB 1051 AGTGGGTAATGCTTATGATG 1110
QY 1010 ATCTGTAATCTGGATCATGTTTC 1069
DB 1111 GPCCTGGCTTCTAGATCATGTT 1170
QY 1070 TAGACTGCTTTAAGGTATGAC 1129
DB 1171 CAGGCTTCATGAGGTATGATG 1230
QY 1130 ATTTGCTAGTACAGTCGACCC 1189
DB 1231 CTTTGTGAGTACAGTGGATCCA 1290
QY 1190 ATCCATTTCTTCTGTTGATTAT 1249
DB 1291 GTCCATTTCTTCTGTTGATTAT 1350
QY 1250 AGCTTCAATGCTGACTTTTGAC 1309
DB 1351 CTCACCTATGCTGACTTTTGAG 1410
QY 1310 ACCATTTCTGAGAGATGATGAC 1369


```
XX FH Key Location/Qualifiers
XX FT CDS 179..3397
XX FT /*cag= a
XX FT /product= "cellulose synthase"
XX FT /note= "no termination codon given"
XX PN WO200009706-A2.
XX XX
XX PD 24-FEB-2000.
XX XX
XX PF 16-AUG-1999; 99WO-US18760.
XX XX
XX PR 17-AUG-1998; 98US-0096822.
XX XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Dhuuga KS, Helentjaris TG, Bowen BA, Wang X;
XX PT WPI; 2000-224343/19.
XX DR P-PSDB; AAY84117.
XX DR
XX PT New genes which encode maize cellulose synthase polypeptides in plants
XX PT useful for modulating the expression of cellulose synthase in plants
XX PT and to produce transgenic plants expressing the novel protein -
XX XX
XX PS Claim 1; Page 160-165; 119pp; English.
XX CC
XX CC The present sequence encodes a maize cellulose synthase polypeptide.
XX CC The cellulose synthase can be used for the improvement of stalk quality
XX CC for improved stand or silage. It also provides an increased concentration
XX CC of cellulose in the pericarp, hardening the kernel and improving its
XX CC handling ability. The sequences are used to produce transgenic plants
XX CC and seeds expressing the cellulose synthase. The polynucleotide is
XX CC used for modulating, preferably increasing, the level of the synthase
XX CC in a plant cell. The plants are preferably monocots. The polynucleotide
XX CC is also used as a probe or primer in the detection quantitation or
XX CC isolation of gene transcripts. The probes are useful in detecting
XX CC deficiencies in the level of mRNA in screenings for desired transgenic
XX CC plant, for detecting mutations in the gene, for monitoring upregulation
XX CC of expression or changes in enzyme activity in screening assays of
XX CC compounds, for detection of any number of allelic variants of the gene,
XX CC or for use as molecular markers in plant breeding programs. The
XX CC isolated nucleic acids of the present invention can also be used for
XX CC recombinant expression of their encoded polypeptides or for use as
XX CC immunogens in the preparation and/or screening of antibodies. The
XX CC proteins can be employed in assays for enzyme agonists or antagonists
XX CC of enzyme function or for use of immunogens or antigens to obtain
XX CC antibodies specifically immunoreactive with a protein.
XX XX
XX SQ Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;

Query Match 38.8%; Score 1408.4; DB 21; Length 3725;
Best Local Similarity 66.8%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;

QY 53 CCGGGCCATGAGCGGACCGGACGCCCTGAAGTCGGGAGGACGGGCGGGGAGCT 112
DB 238 CCGCCAGGAGGGGACCGGCTGTCCCGGCTAAGCCACGAAGAGTGCGAATGGCAGGT 297
QY 113 GTGCCAGATCTGGCCCGACGGCTGGGCACACAGTTCGGACGGCAGCTCTCACCGCCTG 172
DB 298 CTGCCAGATCTGTGGCGACACTGTGTGGCTTTCAGCCACTGTGTGTGTGCTGCTG 357
QY 173 CGACGCTGCGGCTTCGCGGTCTGCCGCCCTGTACGACGAGGCGCAAGGAGGCGAC 232
DB 358 CAATGAGTGTGCCTTCCTCTCTCGCGCCCTTGTATGACTACGAGGCGCAAGGAGGAA 417
QY 233 CGAGGCGCTGCTCAGTGCAGGACCAAGTACAAAGCGCCACAGAGGAGGCGCCAGCTCG 292
DB 418 CCAATGTGCGCCCTCAGTGCAGGACTAGATACAAAGAGACAGAAAGGTAGCCCTCGAGTCA 477
QY 293 CGGGGAGGAGGCGACGACACTGATGCCGATGATGGTAGTGACTTCAACTACCCCTGCATC 352
```

Db	1471	AATAGATTACCTGAAGGACAAAATTCAAACCTTCATTTTGTAAAGAAAGACGACCAATGAA	1530
Qy	1430	GAGAGAAATATGAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAGGCATTTGAAAGT	1489
Db	1531	GAGAGATATGAAGAAATTCAAATTAAGAAATCAATGCCCTTGTGCCAAAGCACAGAAAGT	1590
Qy	1490	CCCCGAGGAAGGATGATCATGCAAGATGGCACACACATGGGCCGAGGAACAATACCAAGGA	1549
Db	1591	GCCTGAAGAGGGGTGCACCATGGCTGATGGAACTGCTTGGCTGGGAATACCCCTAGGA	1650
Qy	1550	TCATCCTGGAATCATTACAGGTTTCCCTTGGTGCACAGTGGTGGCCCTGCATACTCAGGGTAA	1609
Db	1651	CCATCTGGCATGATTCAGGTGTTCTGGGCACAGTGGTGGGCTTGACACTGATGGAAA	1710
Qy	1610	TGAGCTCCCCCGTTTAGTTTATGTCTCGTGAAGCGTCCCTGGGTTCAGACACCACAA	1669
Db	1711	TGAATTACCACGCTTGTCTATGTCTCGTGAAGAGACAGGCTTTCAGCATCACAA	1770
Qy	1670	GAAGGCTGGTCCCATGAATGCCCTTGTTCGTGCTCAGCTGCTCCTTACTTAATGGACAATA	1729
Db	1771	GAAGGCTGGTGCAAATGAATGCACTGATTCGTGTATCTGCTGTGTGCAAAATGGTGCCTA	1830
Qy	1730	CATGTTGAATCTTGATTTGTGATCACTACATCAACAACAGCAAGGCTGTCGGAGAAAGCTAT	1789
Db	1831	TCITTCATGCTGGAATGTGACCATTACTCAATAGCAGCAAAAGCTCTTAGAGAAGCAAT	1890
Qy	1790	GTGCTTCCTTAATGGATCCAAACCTTAGTCCGCCGAAGCTGTGTTATGTGCAGTTCGCCAAG	1849
Db	1891	GTGCTTCATGATGGATCCAGCTCTAGGAAGGAAAACCTTGTATGATAAATTTCCACAAG	1950
Qy	1850	GTTTGATGGGATGATAGGAATGATCGATATGCAACAACAGGAACACTGTCTTTTTTGTATAT	1909
Db	1951	ATTGATGGCATTTGACTTGCACGATCGATATGCTTAATAGGAACATAGTCTCTTTGTATAT	2010
Qy	1910	TAACTTGAGGGCGCTTGACGGCATTCAGGACCAAGCACGTTTATGTGGCACTGGTGTGTTTT	1969
Db	2011	CAACATGAAAGGCTAGATGGCATTCAGGGTCCAGCTCTATGTGGGACACAGGATGCTGTTT	2070
Qy	1970	CAACAGAACAGCTATCTATGGTTTATGAGCC---CCCAATTAAAGCGGAAGACCGAGTTT	2026
Db	2071	CAATAGGCAGCGTTGTATGGATATATGTCCTGTTTGTACTGAAGCTGATCTGGAACCTAA	2130
Qy	2027	CTTGGCATCACTATGTGGGGCAAGAGAAGCAAGCAAGTCAAGAAAGAGAGCTCAGA	2086
Db	2131	CATTGTTGTTAAAGAGTGTGTGGTAGAAGGAAGAGAAGAACAAAGAGTTATATGGATAG	2190
Qy	2087	TAAGAAAAGCTGCAACAAGCATGTGCACAGTTCTCTCCAGTATTTCATCTCGAAGACAT	2146
Db	2191	TCAAAGCCGATATTAGAAGAACAGAATCTTCAGCTCCCATCTTTAAATGGGAAGACAT	2250
Qy	2147	AGAGGAGGGGTGTTCAAGGTGCTGGGTTCATGATGAGAAATCAGTCTCTCATGTCTCAAAAT	2206
Db	2251	CGAGGAGGGTATTGAAGGT-----TATCAGGATGAAAGGTCAGTGCTTATGTCCCAGAG	2304
Qy	2207	GAGCTTAGAGAAGAGATTTGGCCAGCTCAGCAGCATTTGTTGCCCTCCACTCTGATGGATA	2266
Db	2305	GAATTTGGAGAAACGCTTTGGTCAGTCTCCAAATCTTCATTGTCATCCACTTTTATGACTCA	2364
Qy	2267	TGGTGGTGTTCCTTCAGTCTCCACTCCAGAAATCTCTTTGAAAGAAGCATTCCTCATGTCAAT	2326
Db	2365	AGGTGGCATACCACCTTCAACAACCCAGCTTCTCTACTGAAGGAAGCATATCCATGTTAT	2424
Qy	2327	AAGTTGTGGCTATGAGGACAAGTCTGAATGGGGAACTGAGATTGGTTGGATCTATGATCAT	2386
Db	2425	CAGCTGTGGGTACGAGGACAAAACTGAATGGGSAAGAGATTTGGCTGGATCTATGGTTC	2484
Qy	2387	TGTCACAGAGATATCTTACTCGATTCAAGATGCAAGCAAGAGGCTGGCGTTCAGTCTA	2446
Db	2485	AGTTACAGAGGATATCTGACTGGGTTTAAATGATGCAAGAGGCTGGCAATCAATCTA	2544
Qy	2447	TTGCATGCCCAAGCGCCAGCTTTCGAAGGGATCTGCCCCATCAATCTTTCAGATCGCTCT	2506
Db	2545	CTGCATGCCACGAGACTTTTTCAGGGTCTCGACCAATCAATCTTCTGTATCGTCT	2604

RESULT 12	
AAA67114	
ID	AAA67114 standard; DNA; 3851 BP.
XX	
AC	AAA67114;
XX	
DT	31-OCT-2000 (first entry)
XX	
DE	Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.
XX	
KW	Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW	plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW	transgenic plant; ds.
XX	
OS	Pinus radiata.
XX	
PN	WO2000022092-A2.
XX	
PD	20-APR-2000.

XX 08-OCT-1999; 99WO-NZ00169.
XX 13-OCT-1998; 98US-0170862.
XX 11-AUG-1999; 99US-0148426.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX (PLET-) FLETCHER CHALLENGE FORESTS LTD.
XX Bloksberg LN;
XX WPI; 2000-339328/29.
XX P-PSDB; AAB16307.
XX New genes encoding proteins involved in a plant polysaccharide
XX biosynthetic pathway, useful for modulating or altering the
XX polysaccharide content, composition or structure of the plant -
XX Claim 1; Page 71-72; 301pp; English.
XX The present invention describes isolated polynucleotides (PN) comprising
XX a sequence selected from one of 835 nucleotide sequences given in
XX AAA67073 to AAA67907, their (reverse) complements, sequences producing
XX an Expectation (E) value of 0.01 or less compared to the 835 sequences,
XX sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
XX 835 sequences or sequences that are degenerately equivalent or allelic
XX to the 835 sequences. The polynucleotides are used to modify the
XX activity of a polypeptide involved in a polysaccharide biosynthetic
XX pathway in the plant. They are especially used to modulate or alter the
XX polysaccharide content, composition or structure of the plant. AAB16268
XX to AAB16340 are proteins encoded by some of the polynucleotide sequence
XX given in the present invention.
XX Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 other;
Query Match 38.1%; Score 1382; DB 21; Length 3851;
Best Local Similarity 66.4%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 1040; Indels 46; Gaps 9;
QY 75 GAGCGCCCTGAAGTCCGGGAGGACGCGGGGAGGTGCGAGAGTGGCCGCGAGCGG 134
DB 222 GAGGCGCGAAGCCTTTGACACGCTTGAGTGGCCAGCTGCGAGATTTGTGGCGAGAC 281
QY 135 CTGGGACACACGTGGAGCGGACGCTTTCACCGCGCTGGAGCGTGGCGGCTCCCGGTC 194
DB 282 GTGCGGCTTAACACAGACGCGAGCTGTTCGTTGCTGTATGAGTGGGGTTTCTCTGTC 341
QY 195 TGGCGCCCTGTACGAGCAGGCGCAAGGAGGACCCAGGCGCTGCTCCAGTGCAAG 254
DB 342 TGTGCGCGCTGTAGTACGAGACGAGACGAGAGGAAATCAGTGTGCGCGAGTGCAAT 401
QY 255 ACCAAGTACAGCGCCACAGAGGAGGAGCCAGCGATCCCGGGGAGGAGGCGAGCAGACT 314
DB 402 ACTCGTTACAAAGGCTCAAAAGGAGTGCA-----CGGGTGCAAGGTGACGATGAA 455
QY 315 GATGCGCGATGAGTAGTCACTTCACTACCTGATCTGGCACTGAGGACGAGGACAG 374
DB 456 GAAGACGCTTGAACATAGACATGAATTTAATGTGGAGACTCAGCAAGAAACAGGACG 515
QY 375 AAGATTGCTGACAGGATGGCGAGCTGGCGCATGAACACCGGGGCGAGTGCGCAATGTTGGC 434
DB 516 CAGATCAGCGAGCGATGCTCCACGAGCGCATGAGCTATGCGGAGGTCGCGACGACGAA 575
QY 435 CACCCCAAGTATG-ACAGTGGGAGATCGGCTCTCCCA-AGTATGACAGTGGAGATGCC 492
DB 576 AATTGCGAGATTGCTCATATCCAGAGCTTCTCCGCGAGATTCTCTGTACTTTGCAACGGC 635
QY 493 CTAGGGGATACGTCCTTCACTCAGTCACCAACAGCGAGATGTGAGGAGAAATCCCTGGAGCTT 552
DB 636 CACTCGGTTGTAGTGGGAGATTCACAGCTCATACTAGCGAGACACCAATG----- 689
QY 553 CGCCTGATCATCATGATGTCCCTTACCGGGAACATACAGACGCTGCGTTTCCT 612

DB 690 ---CTTGCCAAACCTTGAATGC
QY 613 ATGTGAATCATTCACCAATCC
DB 747 AGGATCATCATGATGATCAACA
QY 673 GGAAGAGAGAGTTGATGGCTG
DB 805 GGAAGGAGCGAGCGATGGTTA
QY 733 ATGGACAAGCATTTGCTCCCTC
DB 865 CGGAAGGGAGATATCAATAA
QY 793 AATACACATGGAAGAGGCTTT
DB 925 ATCCGATATG-----C
QY 853 TCCCATTTGCTTCTCCAAAT
DB 976 TGCAATCTTCAAGCAAAAT
QY 913 TTCTAAGCATTTCTTCGCACTA
DB 1036 TGTGGGTATTTTCTCCGCTA
QY 973 GGCTTTTATCTGTATATGAG
DB 1096 GGGCACTTCTATCGTTTGGAL
QY 1033 CGAAGTGGTTTCAATCAACCG
DB 1156 CCAAGTGGTTGCTATCAGTCG
QY 1093 GAGNAGGTGAACCGTCTCAGTT
DB 1216 GAGAAGCGAACCATCAATGCT
QY 1153 TGAAGGAGCCACCTATCGTCAC
DB 1276 TGAAGGAGCCCTCTTGGTTAC
QY 1213 CCGTGGACAGGCTCTCTGCTAT
DB 1336 CTGTAGACAATGTCTCTGTTA
QY 1273 CATTTGGCTGAGACTTCAGAGTT
DB 1396 CTCTCTGAGACTCAGAAAT
QY 1333 TTGAACCCAGAGCTCCGAGTTT
DB 1456 TTGAGCCTCGCTCCGAAATC
QY 1393 TCCAGCCTTCATTTGTTAAAGC
DB 1516 TTCAACCCACTTTGCAAGAG
QY 1453 TCAGGATAAATGCCCTAGTTCT
DB 1576 TGCGCATCAATGTTGTTGCA
QY 1513 AAGATGGCACACCATGGCCAGGA
DB 1636 AAGACGGTAGCGCTTGGCCTGGT
QY 1573 TCCTTGTCTACAGTGGTGGCCTT
DB 1696 TCTTGGGTACAGTGGCGGCTC
QY 1633 TGTCTGTGAAAGCGTCTCTGGG
DB 1756 TTTCTCGTAGAAGAGACCTGGT

QY 1693 TTGTTCTGCTCTCAGCTCTCTTACTAATGAGCAATATACATGTTGAATCTTCTGATGTGATC 1752
Db 1816 TGGTTGGGGTTTCTGCTGTCTGCTACCAATGTCTCCATTTTATCTGATCTGAATTTGTGATC 1875
QY 1753 ACTACATCAACAACAGCAGGCTGTCGAGAGAGCTATGCTTCTCCTAATGATGATCAAAACC 1812
Db 1876 ACTACATTAACANTAGCAGGCAATCAGGAAAGGATGTGCTTTATGATGATGATCTCTCAGG 1935
QY 1813 TAGGTCCCGCAAGTCTGTTATGTGAGTTCACCAAAAGTTTGTATGGGATTCATAGGAATG 1872
Db 1936 TTGGGAGAAAGTCTGTTATGTGATGTCCTCAGAGATTCGATGCTATGATGCAATG 1995
QY 1873 ATCGATATGCAACAGGACACTGCTTTTGTGATTAATTAAGTTGAGGGCCTTGACGGCA 1932
Db 1996 ACCGTTACGCCAATGAAACCGTATTCTTTGATATCAACATGAAGGCTGCGATGAA 2055
QY 1933 TTCAAGGACCACTTATGTGGAACTGGTTGTGTTTCAACAGAACAGCTATCTATGCTT 1992
Db 2056 TTCAGGGCCTGTATATGTGGAACTGGATGCATGTTCAAGAGACAAGCTCTATATGGGT 2115
QY 1993 ATGAGCCCCAATTAAGCGGAAGAGCCAGTTTCTTGGCATTCATATGTGGGGGCAAGA 2052
Db 2116 ATGGGCTCCCAAGGCCCAAAACGTCCTCAAGATGGTGACCTGTGATTGT-----C 2166
QY 2053 AGAAGCAAGCAAGTCAAGAAAGAGCTCAGATAAGAAAAGTCGAACAGCATGTGG 2112
Db 2167 TCCCTTGTGGGCTCTCTGAAGAGTCTCCGAGAAAATAGTAGCAGAAAAGTGGAG 2226
QY 2113 ACAGTTCTCTTCCAGTATTCAATCTCGAAGACATAGAGAGGGTGTGAAAGTGTGGGT 2172
Db 2227 GAATCCAGCTCCGCGCTACAATCTGGACGGGATCGAGGAAGGATGAGAAGT-----T 2280
QY 2173 TTGATCATGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAGAGAGATTTGGCCAGT 2232
Db 2281 ATGATCAGGAAAGAGCATTTGTCATGAGCCAACTAGACTTCGAGAAGAGTTTGGCCAGT 2340
QY 2233 CAGCAGCATTTGTGCTCCACTCTGATGGAATATGGTGGTGTCTCTCAGTCTCCACTTC 2292
Db 2341 CTTCAGCTTTGTTCATCCACTCTGATGGAGAAATGGTGGTTCGCGCAACAGCAATC 2400
QY 2293 CAGAATCTCTTTGAAGAAGCTATCCATGTCTAATAGTTGTGGCTATAGGACAAGTCG 2352
Db 2401 CAGCTGAATTTGTAAGGAGGCTATTTCATGTCTATCAGCTGTGGATATGAAGACAAACGG 2460
QY 2353 AATGGGAAGTGAATGTTGGATCTATGGATCTCTCAGAGATATTTACTTGGAT 2412
Db 2461 AATGGGAAAGAGCTTGGATGGATCTATGGATCAGTCAAGAGGACATTCGACTGGAT 2520
QY 2413 TCAAGATGCAAGCAGGCTGGCGTTCACTTATTCATGCGTCCAGCGCCAGCTTTCA 2472
Db 2521 TCAAGATGCACACTCGAGGCTGGCGTCCATTTACTGTATGCCAAACGAGCAGCATCA 2580
QY 2473 AGGATCTGCCCCCAATCAATCTTTTCAGATCGCTGAACCAAGTGTGCGGTGGGCTCTCG 2532
Db 2581 AAGGGTGTCTCCAATCAATCTATCAGACGGTTTGAACAGGTTGTGCGTTGGGCTTTGG 2640
QY 2533 GTTCTGTGAAATCTTTTCAGCGCGCATTTGCCCTTATGATGCTACGGAGGCGGC 2592
Db 2641 GATCAGTAGAAATTTTCATGAGCAGACATTTGCCCAATCTGTAAGCTATGGGGAGGTC 2700
QY 2593 TCAAGTTCTCTGGAGAGTTCGCTTACATCAACACCACCTTTTACCACCTAACCTCTCTCC 2652
Db 2701 TGAATGCTTGAAGATTTGCCCTATATCAACACCATTGTCTATCCATTCACCTCTCTTC 2760
QY 2653 CGCTTCTAGTCTATGTATATGCTGCTATCTGCTGCTCACTGGAAGTTTCATCATGC 2712
Db 2761 CACTCATTCGCTATGACACTTCCAGCCGCTCAGTTTGTCTCACTGGCAAAATTTGTGATCC 2820
QY 2713 CAGAGATTAGCAACTGGCCAGTATCTGGTTTCATTCGGCTCTCTTCAATTTTCGCCA 2772
Db 2821 CTCAGATCAGTACTTTTGCAAGTCTATTTTAAATAGTCTCTTTTCATCTCAATTTTGCCA 2880

QY 2773 CTGSTATCCTTGAGTGGAGTGGTGTGTGGCATTTGAGTGGTGGAGGAATGAAC 2832
Db 2881 CTGSTATCTGAAATGAGTGGAGTGGAGTGGATGAGCATTTGAAGATGGTGGCGAAATGAAC 2940
QY 2833 AGTTCTGGGTCAATGGAGGTATCTTGACATCTGTTTGGCGTCTTTTCAGGGTCTTCTGA 2892
Db 2941 AGTTCTGGGTATTTGGAGGGTTCCTGCACATTTTTCGAGTATTTTCAAGGCTCTGCTCA 3000
QY 2893 AGGTCTTGGCGGTATCGACACCAACTTCACTGTCACTCAAGGCTTAATGAGCAAGAAG 2952
Db 3001 AGGTACTGGCAGCATTTGACAAATTTACAGTCACTGCCAAGGC---ATCAGATGACG 3057
QY 2953 GCGACTTTGTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCTCCCGACGACCA 3012
Db 3058 GTGAGTTTGGGAACTGTATGCATTTCAATGGACCACTCTCATTTCTCTTACAACCC 3117
QY 3013 TTTTGTGATCAATTAACATGGTTGGTGTGCTTGGTGGCACTCTTACGCCATCAACAGTGGT 3072
Db 3118 TGCTTGTCAACAACCTTTTGGGGTGTGTTGGCGTAGCAGATGCAATCAACAATGGAT 3177
QY 3073 ACCAATCATGGGCGCGCTCTTTGGGAAGCTCTTCTTGTCTTCTGGGTGATTTGTTCCACT 3132
Db 3178 TTCAGTATGGGTGCTCTCTTGGGTAAAGCTTTTCTTTTGTGATCTGGGTCAATTTGTCACC 3237
QY 3133 TATACCCATTCCTCAAGGGTCTTATGGGCGAGGCAAAACCCGACACACCAGCATTTGTCAFCG 3192
Db 3238 TGTATCTTCTCAAGGGTCTCATGGCAGGCAAGAACCCACCACCTCGTGGTTA 3297
QY 3193 TCTGGGCTGCTCTCTGCTTCTATCTTCTTCTTCTGCTGGTGTGGTTCGTTGATCCATTCA 3252
Db 3298 TTTGGTCAATTTCTCTGCTGCACTGTTTCTCTCTTTTCTGGGTAAAGAAATTTGATCCCTTCT 3357
QY 3253 CTACCGCTCTCGCTGGCCAAATATCAAACTGTGGCATCAACTGCT 3300
Db 3358 TGAGTAAGTTAAAGGCCACAGATACTAAACAATGTGSCATCAACTGCT 3405

RESULT 13
AAZ99515
ID AAZ99515 standard; DNA; 3753 BP.
XX AAZ99515;
XX
XX
XX 03-JUL-2000 (first entry)
XX DNA encoding a maize cellulose synthase.
XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX Zea mays.
FH Key Location/Qualifiers
FT CDS 184..3405
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 916..918, aa: Xaa)
FT /note= "no termination codon given; Xaa is an
FT unspecified amino acid"
XX WO200009706-A2.
PN
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US18760.
XX
XX 17-AUG-1998; 98US-0096822.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
PI
XX WPI; 2000-224343/19.

Db 1779 GAAGCGTGGTCAATGAATGGCGTGAATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTA 1838
Qy 1730 CATGTTGAATCTGATGATGATCAGTACATCAACAACAGCAAGCGTGTCCGAGAACTAT 1789
Db 1839 TCTTCTCAATGTGGATTGCGCAACATTAATCAATAGCAGCAAAAGCTTTAGAGAACAAT 1898
Qy 1790 GTGCTTCTTAATGATCAACAACTAGTCCGCAAGTCTGTATGTGCAGTTCCCAACAAG 1849
Db 1899 GTGCTTCATGATGATCGCGCTCTAGGAAGAAACTGTGTATGTCAATTTCCACAGAG 1958
Qy 1850 GTTTGATGGATGATAGGAATGATCGATATGCAACAGCAAGCACTGTCTTTTGTATAT 1909
Db 1959 ATTTGATGGCATGACATTGCGACGATGATATGCTAATCGGAACATAGTTTCTTTGATAT 2018
Qy 1910 TAACCTTGAGGGCCCTTGACGGCATTCAGGACAGTTTATGTGGGACGTGGTGTGTTTT 1969
Db 2019 CAACATGAAGGCTCTGATGGCAATTCAGGTCAGTTTACGTGGGACAGGATGCTGTTT 2078
Qy 1970 CAACAGAACAGCTATCTATGTTTATGAGCC---CCCAATTAAGCGCAAGAGCCAGGTTT 2026
Db 2079 CAATAGACAGGCTTTGTATGATACGATCCTGTTTCACTGAAGCTGATCTGGAGCCAA 2138
Qy 2027 CTTGGCATCTATGTGGGCAAGAAGAGCAAGCAAGTCAAGAAAGAGAGCTCAGA 2086
Db 2139 CATTTGTTAATGAAGAGTCTGTGTAGAGGAAGAAAGAAAGAGTTATATGGATAG 2198
Qy 2087 TAAGAAAAGTCGAACAGCATGCGACAGTCTGTGCCAGTATCAATCTCCAAGACAT 2146
Db 2199 TCAAGCGGTATTATGAAGAGACAGAAATCTTCAGCTCCCATCTTCAATATGGAAGACAT 2258
Qy 2147 AGAGGAGGTTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAAT 2206
Db 2259 CGAAGAGGTTATTGAAGT-----TACGAGGATGAAGGTCAAGTCTATGTCCAGAG 2312
Qy 2207 GAGCTTAGAAGAGATTTGGCCAGTCAGCAGACATTTGTTGGCTCCACTCTGATGAATA 2266
Db 2313 GAAATTTGGAAGACGCTTTGGTCTCTCTATTTTTCGATCCACCTTTATGACACA 2372
Qy 2267 TGGTGGTCTTCCCTCAGTCCCTCACTCCAGATCTCTTTTGAAGAGACTATCCATGTCAT 2326
Db 2373 AGTGGGATACCACTTCAACAAACCAGCTTCTACTAAGAGAGCTATCCATGTCAT 2432
Qy 2327 AAGTTGTGGCTTAGGACAACTGTAATGGGAACCTGAGATTGGTTGGATCTATGGATC 2386
Db 2433 CAGTTGTGATATGAGCAAACTGAATGGGAAAAGAGATTGGCTGGATCTATGGTTC 2492
Qy 2387 TGTACAGAGATATTCTTACTGATTCAGATGACGACGCAAGAGCTGGCGCTTCAGTCTA 2446
Db 2493 AGTAACGGAGGATATTCTGACTGGGTTTAAATGCAATGCAAGGGGCTGGCAATCAATCTA 2552
Qy 2447 TTGCATGCCAAGCGCCAGCTTTCAAGGATCTGCCCATCAATCTTTTCAGATCGTCT 2506
Db 2553 CTGCATGCCACACAGCTGTTTCAAGGTTCTGCAACCAATCAATCTTTCCGATCGTCT 2612
Qy 2507 GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTGTTGAAATCTTTTTCAGCGCGGATTTGCC 2566
Db 2613 TAATCAGTGTCTCGTGGGCTCTTGGCTCAGTGAATCTCTGTTAGTAGACATTGTCC 2672
Qy 2567 CTTATGTTATGGCTACGAGGCGCCCTCAAGTTCCTCGAGAGATTCGCTTACATCAACAC 2626
Db 2673 TATCTGTATGGTTTACATGACGATTTGAAGCTTTTGGAGAGGCTGGCTTTACATCAACAC 2732
Qy 2627 CACCAATTTACCCTTACCTCTCCGCTCTAGTCTATTTATGATATTTGCTCTATCTG 2686
Db 2733 TATTGTATATCCAATCAATCCATTCGCTTATTTGCTTATTTGCTTATTTGCTTATCTG 2792
Qy 2687 TCTGCTCAGTGGAAAGTTCAATCATGCGACAGATTAGCAACTTGGCCAGTATCTGGTTTCAT 2746
Db 2793 CTCTCTTACCAATAAATTTATCATCTCTGAGATTAGCAATATGCTGGGATGTTCTTCAT 2852
Qy 2747 TGGCGCTCTCTTTCAATTTTCGCCACTGGTATCTCTTTGAGATGAGGTGGGTGGTGG 2806
Db 2853 TCTTCTTTTCCGCTCCATTTTGGCCACTGGTATATTTGGAGCTTAGATGGAGTGGTGTGG 2912

Qy 2807 CATTCAGAGTGGTGGAGAAATGAAACATTTCTGGGTCAATTTGAGGATATCTCTGCACATCT 2866
Db 2913 CATTTGAAGATTTGGTGGAGAAATGAGCACTTTTGGGTTATTTGGTGGCACTCTGCCCATCT 2972
Qy 2867 GTTTGCCGTCTTTTCAAGGTCTTCTGAAGGTCTTCCGCGTATCGACACCAACTTCACGTG 2926
Db 2973 CTTCCAGTGTTCAGAGGCTCTGCTGAAAGTGTGGCTGGGATGATACCAACTTCACAGT 3032
Qy 2927 CACCTCAAAGGCTAATAGCAAGAGGCGACTTTGGCTGAGCTCTACATGTTTCAAGTGGAC 2986
Db 3033 TACCTCAAAGGCACTCTGATGAGGATGGGACTTTTCTGCTGAGCTATATGTTTCAAGTGGAC 3092
Qy 2987 GAGCGTCTCTCACTCCCTCCGACGACCATTTTGTATCATTAACATGTTGGTGTGCTGTGG 3046
Db 3093 CAGTTGCTCATCTCTCCGACCACTGTTCTTGTCAATTAACCTGTGCGAATGGTGGCAGG 3152
Qy 3047 CACCTCTACGCCATCAACAGTGGTTACCAATCATGSGGGCGGCTCTTTTGGAAAGCTCTT 3106
Db 3153 AATTTCTTATGCAATTAACAGTGGCTACCAATCTCTGGGTCCTGGCTCTTTGGAAAGCTCTT 3212
Qy 3107 CTTTGGCTTCTGGGTGATGTTCACTTATACCCATTTCTCAAGGTCTTATGGGCGAGCA 3166
Db 3213 CTTCTCGATCTGGGTGATCCTCCATCTCTACCCCTTCTCAAGGTCTCATGSGGAGGCA 3272
Qy 3167 AAACCGCACACGACGATGTCATGCTCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 3226
Db 3273 GAACCGCACACCAACAATGCTCATTTGCTGGTCCATCTCTTTCGATCTATCTCTCTCTCT 3332
Qy 3227 GCTGTGGGTTCTGTGTGATPCCATTCACATACC 3258
Db 3333 GCTGTGGGTTGAAGATCGATCCTCTTCATCTCCC 3364

RESULT 14
AAZ99530
ID AAZ99530 standard; DNA; 3753 BP.

XX AAZ99530;

XX 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.

XX Zea mays.

XX Key Location/Qualifiers

XX CDS 184..3405

XX /tag= a

XX /product= "cellulose synthase"

XX /trans_except= (pos: 916..918, aa: Xaa)

XX /note= "no termination codon given; Xaa is an

XX unspecified amino acid"

XX WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US18760.

XX 17-AUG-1998; 98US-0096822.

XX (PION-) PIONEER HI-BRED INT INC.

XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

XX P-PSDB; AAY84120.

XX New genes which encode maize cellulose synthase polypeptides in plants

Db 1839 TCTTCAATGTCGATGGACCATTAATACACGAAGCTCTTAGAGAGCAAT 1898
Qy 1790 GTGCTTCTTAATGATCCAAACCTAGGTCCGCAAGTCTGTTATGTGACGTTCCCAAAAG 1849
Db 1899 GTGCTTCATGATGATCCGGCTCTAGGAAGGAAACCTGTTATGTACAAATTTCCACAG 1958
Qy 1850 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1909
Db 1959 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2018
Qy 1910 TAACTTGAGGGGCTTGACGGCAATCAAGGACCAAGTTATGTGGGAACCTGGTGTGTTT 1969
Db 2019 CAACATGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2078
Qy 1970 CAACAGAACAGTATCTATGTTTATGAGCC---CCCAATTAAGGGCAAGAACCAAGTTT 2026
Db 2079 CAATAGACAGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2138
Qy 2027 CTTGGCATCACTATGTGGGGCAAGAGGCAAGCAAGTCAAGAAAAGNGCTCAGA 2086
Db 2139 CATTTGTTAAGAGCTGCTGTTGTAAGAGGAAGAAAAGCAAGATATATGATGATG 2198
Qy 2087 TAAGAAAAGTCAAGCAAGCATGTGGACAGTCTGTTCCAGTATTAATCTCGAAGACAT 2146
Db 2199 TCAAGCCGTTATTAAGAGACAGATCTTCAGTCCCACTTCCAATATGAAGACAT 2258
Qy 2147 AGAGGAGGTTGTAAGGTCGCTGGTTGATGATGAGAATCAAGTTCTCATGCTCAAAAT 2206
Db 2259 CGAAGAGGTTATTGAAGT-----TACGAGGATGAAGGTCAAGTCTTATGTCACAG 2312
Qy 2207 GAGCTTAGAAGAGATTTGGCCAGTCAGCAGCATTTGTCCTCCACTCTGATGGAATA 2266
Db 2313 GAAATTGGAGAAACGCTTTGGTCAGTCTCCTATTTCATGTCATCCACCTTTATGACACA 2372
Qy 2267 TGTGTTGTTCTCCAGTCCCTCACTCCAGAACTCTTTTGAAGAGGCTATCATGTCAT 2326
Db 2373 AGTGGCATACACCTTCAACAACCCAGCTCTCTACTAAGGAAGCTATCAATGTCAT 2432
Qy 2327 AAGTTGTGGTATGAGACAAAGTCTGAATGGGAACTAGATTTGTTGGATCTATGATC 2386
Db 2433 CAGTTGTGGTATGAGACAAAATGAATGGGAAAAGATGTCGTGGATCTATGTTTC 2492
Qy 2387 TGTACAGAGATATCTTACTGATTCAGATGACGCAAGAGGCTGGCTTCAGTCTA 2446
Db 2493 AGTAACGGAGATATCTGACTGGTTTAAATGCAATGAAGGGCTGGCAATCAATCTA 2552
Qy 2447 TTGCATGCCCAAGCGCCAGCTTCAAGGATCTGCCCATCAATCTTTTCAGATCGTCT 2506
Db 2553 CTGCATGCCACCGACCTTGTTCAGGGTCTGCAACCAATCAATCTTCCGATCGTCT 2612
Qy 2507 GAACCAAGTCTGGGTGGGCTCTGGTCTGTTGTAATTTCTTTTCAGCGGCAATGGCC 2566
Db 2613 TAATCAGTGTCTGGTGGGCTCTGGGTCAAGTGAATTTCTGCTTAGTAGACATTTGCC 2672
Qy 2567 CTTATGATGATGCTACGAGGCGCTCAAGTTCTCGGAGATTCGCTTACATCAACAC 2626
Db 2673 TATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2732
Qy 2627 CACATTTACCCCACTAACCTCTCTCCGGCTTCTAGTCTATGTTATATGCTGCTATCTG 2686
Db 2733 TATGATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2792
Qy 2687 TCTGCTCACTGGAAGTTATCATCATGACGAGATAGCAACTTGGCCAGTATCTGTTTCAT 2746
Db 2793 CCTCTTACCAATAAATTTATCATCTCTGAGATGACCAATTTATGCTGGGATGTTCTCAT 2852
Qy 2747 TGCCTCTCTCTTCAATTTTCGCACTGCTATCTCTGATGATGATGATGATGATGATGATG 2806
Db 2853 TCTTCTTTTCCCTCCATTTTGGCACTGATATATGAGCTTATGAGTGGTGGTGG 2912
Qy 2807 CATTGACAGTGGTGGAGGAATGAACAGTTCTGGGTCAATGAGGATCTCTGCAATCT 2866

Db 2913 CATTCAGATTTGGTGGAGAAATGAGCAGTTTGGTTATTGGTGGCACTCTGCCATCT 2972
Qy 2867 GTTTGGCGCTTTTCAGGCTCTTCTGAAGGTCCTTTCGCGGTATTCACAACTTCACTGT 2926
Db 2973 CTTCGCACTGTTCCAGGCTCTGCTGAAAGTGTTCGCTGGGATGATACCACTTCAAGT 3032
Qy 2927 CACCTCAAGGCTAATGACGAAGAGGCGACTTTTGTCTGAGCTCTACATGTTCAAGTGGAC 2986
Db 3033 TACCTCAAGGCACTCTGATGAGATGGCGACTTTCCTGAGCTATATGTTCAAGTGGAC 3092
Qy 2987 GAGCGTTCTCATCTCCCTCCGACGACATTTTGTATCAATTAACATGTTGTTGCTGG 3046
Db 3093 CAGTTGCTCATCTCCTCGACCACTGTTCTTGTTCATTAACCTGTCGGAATGTTGCGAG 3152
Qy 3047 CACCTCTACCCCACTCAACAGTGTACCAATCATGAGGGGCGCTCTTTGGGAAGCTCTT 3106
Db 3153 AATTCTTATCCCATTAACAGTGGCTACCAATCTCGGCTCCGCTCTTTGGAAGCTGTT 3212
Qy 3107 CTTTGGCTTCTGGTGTGATTTTCACTTATACCCATTTCTCAAGGCTCTTATGGGCAAGCA 3166
Db 3213 CTTCTGATCTGGGTGATCTCCATCTACCCCTTCTCAAGGCTCTCATGGGAAGCA 3272
Qy 3167 AAACCGCACACGAGATTTTCATCGTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 3226
Db 3273 GAACCGCACACCAACATCGTCAATTTGCTGGTCCATCTCTTTCATCTATCTCTCTCT 3332
Qy 3227 GCTGTGGGTTGCTGTTGATCCATTCATACCC 3258
Db 3333 GCTGTGGTGAAGATCGATCTCTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3364

RESULT 15
AAZ99497
ID AAZ99497 standard; DNA; 3780 BP.
XX
AC AAZ99497;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker; ss.
XX Zea mays.
FH Key Location/Qualifiers
CDS 201..3422
FT /*tag= a
FT /product= "cellulose synthase"
FT /transl_except= (pos: 933..935, aa: Xaa)
FT /note= "no termination codon given; Xaa is an
XX unspecified amino acid"
PN WO200009706-A2.
XX
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
XX
PR 17-AUG-1998; 98US-0096822.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX
DR WPI; 2000-224343/19.
XX
XX P-PSDB; AAY84109.
XX
XX New genes which encode maize cellulose synthase polypeptides in plants
XX useful for modulating the expression of cellulose synthase in plants
XX and to produce transgenic plants expressing the novel protein -
XX

PS Claim 1; Page 97-102; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.

SQ Sequence 3780 BP; 946 A; 817 C; 976 G; 1040 T; 1 other;

Query Match 37.9%; Score 1373; DB 21; Length 3780;
Best Local Similarity 66.3%; Pred. No. 0;
Matches 2130; Conservative 0; Mismatches 986; Indels 96; Gaps 7;

QY 53 CCGCGCCATGGACGGCGGACGCCCTGAAGTCCGGGAGGCACGGGGCGGGACGT 112
DB 260 CCGCGACGACGGCGATGTCCGGGGTCCGGCTAAGCCCAAGAGGTGCGAATGGACAGST 319
QY 113 GTGCGAGATCTCGCGCGGCGGCTGGGACACACGCTTGGAGCGGCGCTTCACCGCCGTG 172
DB 320 CTGCGCAGATTTGGGTGACTCTGTGGGTGTTTCAGCCACCTGGTGATGCTTTGTTGCCNG 379
QY 173 CGAGCTCTCCCGCTTCCCGGCTGCGCGCCCTGCTACGAGCAGCAGCGCAAGGAGGCGAC 232
DB 380 CAATGAGTGTGCTTCCCTGTGTCGCCCGCATGCTATGATGATGAGCGCAAGGAGGGAA 439
QY 233 CCAGGCGCTGCTCCAGTGCAGACCAAGTACAAAGCGCCACAGAGGAGGCCAGCGATCG 292
DB 440 CCAATGCTGCCCGCCAGTGCAGACTAGATACAAAGAGAGAGAGTAGCCCTCGAGTTCA 499
QY 293 CGGGGAGGAAGCGACGACACTGATGCCGATGATGGTAGTACTTCAACTACCTGCAATC 352
DB 500 TGGTGATGAGGATGAGGAAGATGTTGATGACCTGACACATGAATCAACTAC----- 551
QY 353 TGGCAGTGAAGACCAAGACAGAGATTTGCTGACAGGATGCGCAGCTGGCGCATGAACAC 412
DB 552 -----AAGCAAGGC 560
QY 413 CGGGGCGAGTGCATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTTCCAA 472
DB 561 AGTGGGAAGGCCAGAGTGGCACTGCAAG-----GAGATGATGCTGATGCTGCTTCATCT 617
QY 473 GTATCAGAGTGAGAGATCCCTAGGGGATACGTCCTTCAGTCACCAACAGCCAGATGTC 532
DB 618 GCTCCGCATGACCCACATCATCGGATTCACGCCCTGACAAAGCGGTCAA-----CAGATATC 673
QY 533 AGGAGAAATCCCTGGAGCTTCCCTGATCATCATGATGTCCTTACGGGGAACATCAG 592
DB 674 TGGAGAGATTCCTGATGCTTCCCTGACCGTCAATCTATCCGACGTCCCAACATCAGACTA 733
QY 593 CAGAGCTGCTCGTTCCTTATGTAATCATTCACCAATCCGTCAGGGAGTTC---TC 649
DB 734 TGTGATCCAGCGTCCAGTTCCTGTGAGGATTTGGACCCCTCGAAGGACTTGAATTC 793
QY 650 CGGCGAGTATTGGGAATGTTGCTTGGAAAGAGAGAGTGTGATGCTGGAAATCAAGCAGGA 709
DB 794 CTATGGCTTAATAGTGTGACTGGAAAGAAAGATTTGAGAGCTGGAGGTTTAAACAGGA 853

QY 710 CAAGGGTGGGATTCCTCCATGACT 769
DB 854 CAAAATATGATCAAGTGAAT 913
QY 770 TACTGACATCGATGATCACT 829
DB 914 GACTGCTCAAAATGGA----- 955
QY 830 TCGCCAGCGCTCTATCTAGAA 889
DB 956 AGGCTTACCTTTGAGCGGTATG 1015
QY 890 GGTCAATGTTTCCGGTGGT 949
DB 1016 AGTGATCAATCCCGTCTATG 1075
QY 950 TGTGCGTAATGATACCCACTG 1009
DB 1076 AGTGCCTGATGCTTATGATTA 1135
QY 1010 ATCCTGGTACTGGATCAGTTC 1069
DB 1136 GTCTTGGCTTCTAGATCAGTTC 1195
QY 1070 TAGACTGGGTTTAAAGTATGAG 1129
DB 1196 CAGGCTTGCATGAGGTATGAT 1255
QY 1130 ATTTGCTACGTACGTCGACCC 1189
DB 1256 CTTGCTGAGTACAGTGGATCCA 1315
QY 1190 ATCCATTCCTGCTTGTGATTAT 1349
DB 1316 GTCCATTCCTTCTGTGGATTAG 1375
QY 1250 AGCTCAATGCTGACCTTTTGG 1309
DB 1376 TTCAGCTATGCTGACTTTTGG 1435
QY 1310 ACCATTTGTGAAGAAGTATGAC 1369
DB 1436 TCCCTTTTGTGAAGAAGCAAT 1495
QY 1370 AATTGATTACCTGAAAGACAA 1429
DB 1496 ATAGATTAACCTGAAGGACAA 1555
QY 1430 GAGAGAAATGAAGAAATTTAA 1489
DB 1556 GAGGGAGTATGAAGAAATTCAA 1615
QY 1490 CCCCAGGAGGATGGATCATG 1549
DB 1616 GCCTGAAGAGGGTGGACCATG 1675
QY 1550 TCATCTGGAATGATTCAGGTT 1609
DB 1676 CCATCTGCGATGATTCAGGTT 1735
QY 1610 TGAGCTCCCGCTTTAGTTTAT 1669
DB 1736 TGAGTTACCACTGTTGCTAT 1795
QY 1670 GAAGGCTGTGCCATGAATGCC 1729
DB 1796 GAAGGCTGTGCAATGAATGCC 1855
QY 1730 CATGTTGAATCTTGATTTGAT 1789
DB 1856 TCTTCTCAATGTTGATTTGGAC 1915

SGACAAGCATTTGCTCCTCTCTGAAGGTGGGCGAGC 769
AATATCCAGAGGCTAGAGGAGGACATGGAGGG 913
ACAACATGGAAGAGCGTTTACTGAATGATGAAC 829
-----GAANATATGCAAAATGGTTGATGTC 955
CCATTGCTTCTCCAAATAAATCCCTACAGAAT 889
CAATTCTCTCAAACCAAGCTCAACCTTTACCGGGT 1015
TAAGCATCTTCTCGTGCACACTACCGTCTCAAAATCC 949
TGTGCTTCTTCTCCAGTATCGTGTGATCATCC 1075
TTTATCTGTTATATGATGAGATTTGGTTTGGTTT 1009
TAGTATCTGTTATCTGCGAGGCTCTGTTGCCCT 1135
AGTGGTTTCCAATCAACCGGGAGACCTTACCTTGA 1069
AATGGTATCCAATCAACCGTGTGAGACATATCTTGA 1195
AAGTGAAACCGTCTCAGTTGGCTGCTGTTGACAT 1129
AGGAGAGCCATCACAGCTGGCTCCCATGATGT 1255
AGGAGCCACCTATCGTCACCTGCCCAACACTGTCT 1189
AGGAACCTCCACTGATCACAGCCCAACACTGTTTT 1315
TGGACAAGGTCTCTTGTCTATGTTATCTCATGACGG 1349
TTGACAAAGTGTCTATGTTCTCTGATGATGG 1375
TGGCTGAGACTTCAGAGTTTGGCTAGGAAATGGGT 1309
TCTCAGAAACCGGAGAAATTTGCTAGAAAGTGGT 1435
AACCCAGAGCTCCGAGTTTACTTTTTCGCCAGAA 1369
AACCAAGAGCTCCAGAAATTTACTTGTCTCAAAA 1495
AGCTTCAATTTGTTAAAGACCGCGGGCCATGAA 1429
ACCTTCAATTTGTTAAGGAAGAGCGGCATGAA 1555
SGATAAATGCCCTAGTTTCTTAAGGCATTTGAAGT 1489
GAATCAATGCCCTTGTGTCACCAAGCACAGAAAGT 1615
ATGGCACACCATGGCCAGGAGAAACATACCAGGGA 1549
ATGGAACATGATGCCCTGGAATTAATCCTAGGGA 1675
TTGCTCACAGTGTGGCTTTGATCTCAGGCTAA 1609
TGGGACACAGTGTGGCTGCTGACACTGATGGA 1735
TCTGTAAGAAAGCGCTCTGGGTTCCAGCACACAA 1669
TCTGCTCAGCTGCTCTTACTTAATGGACAATA 1729
TCTGCTATCTGCTGCTGACAAATGGTGCCT 1855
ACATCAACACAGACAGGCTGTCCGAGAACTAT 1789
ACTTCAATAGCAGCAAGCTCTTAGAGAACAT 1915

QY 1790 GTGCTTCTTAATGGATCCAAACCTAGGTCGCCAAGCTCTGTTATGTGCGAGTTCCACAAAG 1849
Db 1916 GTGCTTCAATGATGGATCCGGCTCTAGGAAGGAAACCTGTTATGTACAAATTCACACAG 1975
QY 1850 GTTTGATGGGATGATAGGAATGATCGATATGCAAAACAGGAACTGCTCTTTTGTGATAT 1909
Db 1976 ATTTGATGCCATGACTTGCAGATCGATATGCTAATCGGAACATAGTTTCTTTGATAT 2035
QY 1910 TAACTTGAGGGCCCTGAGGCCATTCAGGACCACTTTATGTGGAACTGGTTGTGTTTT 1969
Db 2036 CAACATGAAAGCTCTGGATGGCAATTCAGGGTCCAGTTACGTTGGGAACAGGATGCTGTTT 2095
QY 1970 CAACAGACAGCTATCTATGTTATGAGCC---CCCAATTAAGGCGAAGACCGAGTTT 2026
Db 2096 CAATAGACAGGCTTTGTATGATATGATGCTCTGTTTGAAGTGAAGTATGCTGGAGCCAAA 2155
QY 2027 CTTGGCATCACTATGTGGGGGCAAGAAGGCAAGCAAGTCAAAAGAAAAGGAGCTCAGA 2086
Db 2156 CATTTGTTAAGAGCTGCTGTGGTAGAAGGAAGAAAGAACAGAGTTATATGGATAG 2215
QY 2087 TAAGAAAAAGTCGAACAGCATGTGGACAGTTCTGTTCCAGTATTCATCTCGAAGACAT 2146
Db 2216 TCAAAGCCGTATTATGAAGAGAACAAGATCTTCAGCTCCCATCTTCAATATGGAAGACAT 2275
QY 2147 AGAGGAGGCTGTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCCTCAAT 2206
Db 2276 CGAAGAGGCTATTGAAGGT-----TACGAGGATGAAGGTCAGTGTATGTGCCAGAG 2329
QY 2207 GAGCTTAGAGAGAGATTGGCCAGTCAGCAGCATTTGTTGCTCCACTCTGATCGAATA 2266
Db 2330 GAAATGGAGAAACGCTTTGGTCAGTCCTCTATTTTCATGCTCCACCTTTATGACACA 2389
QY 2267 TGGTGTGTTCTCAGTCTCCACCTCCAGAACTCTTTTGAAGAAGCTATCCATGTGAT 2326
Db 2390 AGGTGCATACCACTTCAACAAACCCAGCTTCTACTAAAGGAAGCTATCCATGTGAT 2449
QY 2327 AGTTGTGCTATGAGGAAAGCTCTGAATGGGAACTGAGATTGGTTGGATCATGGATC 2386
Db 2450 CAGTTGTGATATGAGGACAAACCTGAATGGGAAAGAGATTGGCTGGATCTATGGTTC 2509
QY 2387 TGTACAGAGATATTCTTACTGGAATTCAGATGCACGCAAGAGGCTGGCGTTTCACTGA 2446
Db 2510 AGTAACGGAGGATATTCTGACTGGGTTTAAATGATGATCAAGGGCTGGCAATCAATCA 2569
QY 2447 TTGCATGCCCAAGGCCCAAGCTTTCAGGATCTGCCGCCATCAATCTTTCAAGATCGTCT 2506
Db 2570 CTGCATGCCCAACGACCTTGTTCAGGGTTCTCAGCAATCAATCTTTCCGATCGTCT 2629
QY 2507 GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTGAAATCTTTTCAGCGGCATTTGCC 2566
Db 2630 TAATCAGGTGCTCCGTTGGGCTCTTGGGTGAGTGGAAATTCGCTTAGAGACATTTGCC 2689
QY 2567 CTTATGATGGCTACGGAGGGCCCTCAAGTTCTCTGGAGAGATTGCTTTACATCAACAC 2626
Db 2690 TATCTGATGTTACAATGGAGGATTGAAGCTTTTGGAGAGGCTGGCTTTACATCAACAC 2749
QY 2627 CACCATTTACCCACTAACCTCTCTCCGCTTCTAGTCTATTTGATATATGCTGCTATCTG 2686
Db 2750 TATTGTATATCCAAATCACATCCATTCGGCTTATGCTTATGTTGCTTTCCGCTATCTG 2809
QY 2687 TCTGCTCACTGGAAAGCTTATCATGCGAGATTAGCAACTTGGCAGTATCTGGTTTCAT 2746
Db 2810 CTTCTTACCATAAATTTATCATCTCTGAGATTAGCAATATGCTGGGATGTTCTTCAT 2869
QY 2747 TGGGCTCTTCTTCAATTTTCCCACTGGTATCTCTTGGATCAGGTGGAGTGGTGGTGG 2806
Db 2870 TCTTCTTTTCGCCCTCCATTTTGGCCACTGATATATGGAGCTTAGATGGAGTGGTGG 2929
QY 2807 CATTTAGAGGTGTTGGAGAAAGAACAGTTCTGGGTCTATGGAGGTATCTCTGCACATCT 2866
Db 2930 CATTTGAAGATTGGTGGAGAAATGAGCAGTTTGGGTTATTGGTGGCACCCTGCCATCT 2989
QY 2867 GTTTGCCGCTCTTTCAGGGTCTCTGAAAGGTGCTTGCCGGTATCGACACCACTTCACTGT 2926

Db 2990 CTTCCGAGTGTTCAGGGTCTGCTGAAAGTGTGGCTGGGATGATACCAACTTCACAGT 3049
QY 2927 CACCTCAAAGGCTAATGACGAAGAAGCGAGCTTTGTGTAGCTCTACATGTTCAAGTGGAC 2986
Db 3050 TACCTCAAAGGCATCTGATGAGGATGCGAGCTTTGTGTAGCTATATGTTCAAGTGGAC 3109
QY 2987 GACGCTTCTCATCCCTCCGAGGACCATTTTGCATCAITTAACATGGTTGGTGTGCTGCTGG 3046
Db 3110 CAGTTTGTCTATCTCTCCGACCACCTGTTCTGTTCATTAACCTGGTCGGAATGTTGGCAGG 3169
QY 3047 CACCTCTAGGCCATCAACAGTGGTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTT 3106
Db 3170 AATTTCTTATGGCCATTAACAGTGGCTACCAATCTCGGGTCCGCTCTTTGGAAGCTGTT 3229
QY 3107 CTTTGCCTTCTGGGTGATTTGTCTACTTATACCCCAATTCCTCAAGGGTCTTATGGGCAGGCA 3166
Db 3230 CTTCTCGATCTGGGTGATCTCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGAAGGCA 3289
QY 3167 AAACCGCACACCGAGGATTTGTCATCTGCTGGGCTGCTCTCTCGCTTCTATCTTCTCTT 3226
Db 3290 GAACCGCACACCAACAATCGTCATTTGTCTGGTCCATCTCTTGTGCATCTATCTTCTCTT 3349
QY 3227 GCTGTGGGTTGCTGTGTTGATCCATTCACATACC 3258
Db 3350 GCTGTGGGTGAAGATCGATCTTTCACTCTCC 3381

Search completed: February 25, 2003, 18:32:22

Job time : 730 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 15:04:05 ; Search time 4739 Seconds

(without alignments)
12391.820 Million cell updates/sec

Title: US-09-900-237-29

Perfect score: 3626

Sequence: 1 gcagaggaaccccgctcca.....tcacatttgaggaggtttt 3626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_estl: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2572.2	70.9	3763	11	AV108113 Zea mays
2	2402	66.2	3728	11	AV112236 Zea mays
3	2199	60.6	3696	11	AV110079 Zea mays
4	1373.2	37.9	3898	11	AV110415 Zea mays
5	1320.4	36.4	2872	11	AV104236 Zea mays
6	1319.8	36.4	3788	11	AV103701 Zea mays

7	1315.6	36.3	3783	11	AV104730 Zea mays
8	1267.8	35.0	3897	11	AV103655 Zea mays
9	762.2	21.0	880	14	BM16138 HCl09B12
10	751.6	20.7	916	14	BM16139 HCl05G09
11	750.4	20.7	762	14	BQ802778 WHE2829_H
12	729.4	20.1	870	12	BG368813 HVSME1002
13	689.6	19.0	767	14	BQ578769 WHE0309_E
14	689.4	19.0	710	14	BQ282057 WHE3028_D
15	687.8	19.0	750	12	BF624748 HVSME001
16	684.2	18.9	712	14	BQ620630 TaLR1138D
17	683.8	18.9	978	12	BG321258 Zm04_03F1
18	668.6	18.4	755	12	BF259426 HVSMEF001
19	663	18.3	705	9	AJ433167 AJ433167
20	651.4	18.0	659	13	BJ214460 BJ214460
21	649.6	17.9	678	9	AJ433236 AJ433236
22	644.4	17.8	659	14	BQ620294 TaLR1173A
23	637.6	17.6	676	14	BQ468579 HMO1K12T
24	630	17.4	654	10	AV914933 AV914933
25	628.4	17.3	661	13	BJ277746 BJ277746
26	628	17.3	656	14	BQ620013 TaLR1138D
27	616.6	17.0	627	13	BJ287480 BJ287480
28	615.6	17.0	694	13	BJ282938 BJ282938
29	614.8	17.0	634	13	BJ280838 BJ280838
30	611.2	16.9	625	13	BJ314568 BJ314568
31	609.2	16.8	630	13	BJ317016 BJ317016
32	601.6	16.6	671	9	AL504419 AL504419
33	600.2	16.6	680	10	AV920051 AV920051
34	592.8	16.3	672	13	BJ280905 BJ280905
35	591.6	16.3	643	14	BQ465284 H003C22T
36	588.8	16.2	625	14	BQ471177 HVO1F24T
37	581.4	16.0	603	10	BE403558 WHE0434_A
38	580.6	16.0	615	10	AV932043 AV932043
39	580.4	16.0	669	13	BJ466285 BJ466285
40	578.2	15.9	634	12	BG414971 HVSMEK000
41	577.6	15.9	584	13	BJ233686 BJ233686
42	576	15.9	593	13	BJ320053 BJ320053
43	576	15.9	631	10	AV913277 AV913277
44	571	15.7	583	13	BJ207181 BJ207181
45	571	15.7	681	13	BI959339 HVSME001

ALIGNMENTS

RESULT 1	AV108113	AV108113	3763 bp	mrna	linear	HTC 25-MAY-2002
LOCUS	Zea mays PC0126465 mRNA sequence.					
DEFINITION	Zea mays					
ACCESSION	AV108113					
VERSION	AV108113.1 GI:21211191					
KEYWORDS	HTC.					
SOURCE	Zea mays.					
ORGANISM	Zea mays					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.					
AUTHORS	1 (bases 1 to 3763) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2 (bases 1 to 3763) Coe,E.C.					
AUTHORS	Direct Submission					
TITLE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
JOURNAL	Missouri, Columbia, MO 65211, USA					
FEATURES	Location/Qualifiers					
Source	1. 3763 /organism="Zea mays" /db_xref="MaizeDB:637367" /db_xref="taxon:4577" /clone="PC0126465"					

/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"									
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"									
BASE COUNT	863 a	984 c	1011 g	904 t	1 others				
ORIGIN									
Query Match	70.9%; Score 2572.2; DB 11; Length 3763;								
Best Local Similarity	86.4%; Pred. No. 0;								
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;									
QY	4	CGAGGAACCCGCTCCAGCTCTGTCGTCGCGGGTTGGATCGCTCTCCGCGCGCATGG	63						
Db	282	CCAGGCCCCAGGCTCCAGGCGAGCTCCTCGACGCTTTCTCGCGAGCTCGCTTGCCATGG	341						
QY	64	ACGGCGACGCGACGCCCTGAAGTCCGGGAGGCGACGGCGGGGAGCGTGTGCCAGATCT	123						
Db	342	AGGGCGACGCGACCGCGCTGAAGTCTGGGAGGCGCGTGGCGACAGGTGTGCCAGATCT	401						
QY	124	GGCCGACGGCTGGGCAACACAGCTTGGACGGCGACGCTTTCACCGCCTGCGACGTGCGC	183						
Db	402	CGCGGCGACGGCTGGGCAACACAGCGCGGAGGGGAGCTTTCGCGCCTCGGACGTCTGGC	461						
QY	184	GCTTCCCGTCTGCGCGCCCTGCTACGAGCAGCAGCGCAGGAGGCGACCCAGCCCTGCC	243						
Db	462	GGTTTCGGGTGTCGCGCCCTGCTACGAGTACGAGCGCAAGGACGCGACGCGAGGCGTGC	521						
QY	244	TCCAGTCAAGACCAAGTACAAGCGCCACAGAGGAGGCCACCGCATCCGCGGGAGGAAG	303						
Db	522	CCAGTGTCAAGACCAAGTACAAGCGCCACAAAGGGAGSCCCGCGCATCCGTGGGGAGGA	581						
QY	304	GCGAGCACTGATGCCGATGATGGTAGTACCTTCACTACCTGCTGCTGCGACGTGAGG	363						
Db	582	GAGCGCACTGATGCCGA-----TAGCGACTTCAATTACCTTGGATCTGGCAATGAGG	635						
QY	364	ACCAGAAGCAGAAGATTGCTGACAGGATGCGCAGCTGGCGCATGAACACGGGGGAGTG	423						
Db	636	ACCAGAAGCAGAAGATTGCCGACAGATGCGCAGCTGGCGCATGAACGTTGGGGCAGCG	695						
QY	424	GCATGTTGGCCACCCCAAGTATGACAGTGGCGAGATGGCCCTCTCCAAGTATGACAGTG	483						
Db	696	GGGATGTTGGTCGCCCAAGTATGACAGTGGCGAGATGGGGCTTACCAAGTATGACAGTG	755						
QY	484	GAGAGATCCCTAGGGGATACGCTTCACTACCAACAGCCAGATGTCAAGAGAAATCC	543						
Db	756	CGGAGATCCCTCGGGATACATCCCATGCTACACGCCAGATCTCAGGAGAAATCC	815						
QY	544	CTGGAGCTTCCCTGATCATCATATGATGTCCTTACGGGAAACATCAGCAGACGTCCTC	603						
Db	816	CTGGTGTCTCCCTGACCATCATATGATGTCCTTACGGGAAACATGCGCAAGCTGCTC	875						
QY	604	GCTTCCCTATGTAATCATTCACCAATCCGTCAAGGAGTTCCTCCGCGCATTTGGGA	663						
Db	876	CATTTCCCTATGTAACCATTCGCCAAATCCGTCAAGGAGTTCCTCGTAGCATTTGGGA	935						
QY	664	ATGTTGCCCTGGAAGAGAGAGTGTGATGGCTGGAATAATGAAGCAGGACAAAGGTCGATTC	723						
Db	936	ATGTTGCCCTGGAAGAGAGAGTGTGATGGCTGGAATAATGAAGCAGGACAAAGGTCGATTC	995						
QY	724	CCATGACTAATGGGACAGCATTCCTCCCTGTAAGTCTGGGAGCGTACTGACATCATGATG	783						
Db	996	CCATGACGAATGGCACAAGCATTCCTCCCTGTAGGGTGGGGTGTGGTATTGATGATG	1055						
QY	784	CATCTACTGAATCAACATGAAGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTAT	843						
Db	1056	CATCACTGATTTACAACATGGAGATGCCCTTATGAAAGCAGGAACCTCGACAGCCTCTAT	1115						
QY	844	CTAGAAAAGTCCCCCATTCCTTCCCAAAATAATCCCTACAGAATGGTCATTTGTTCTGC	903						

Db	1116	CTAGGAAAGTTCCACATCTCTT	1175						
Qy	904	GGTTGGTGTCTTAAGCATCTT	963						
Db	1176	GATTGATGTCTTAAGCATCTT	1235						
Qy	964	ACCCACTGTGGCTTTTATCTG	1023						
Db	1236	ACCATATAGGCTTCTATCTG	1295						
Qy	1024	ATCAGTTCGCCAAGTGGTTTC	1083						
Db	1296	ATCAGTTCGCCAAGTGGTTTC	1355						
Qy	1084	GGTATGACCCAGAGGTGAAC	1143						
Db	1356	GGTATGACCCAGAGGTGAAC	1415						
Qy	1144	TCGACCCCTTTGAAGAGGCCAC	1203						
Db	1416	TCGACCCCAATGAAGAGCCCTC	1475						
Qy	1204	TTGATATCCCGTGGACAAGGT	1263						
Db	1476	TGGATTACCCGTGGATAAGGT	1535						
Qy	1264	CTTTTGACGCAATTGGCTGAGAC	1323						
Db	1536	CAITTGATGCACTAGCTGAGAC	1595						
Qy	1324	AGTATGACATTTGAACCCAGAGC	1383						
Db	1596	AGTACAACTTTGAACCTTAGAGC	1655						
Qy	1384	AAGACAAAGTCCAGCCCTTCAT	1443						
Db	1656	AGCAAAAGTGCACCTTCAT	1715						
Qy	1444	AATTTAAATCAGGATAAATGC	1503						
Db	1716	AATTTAAAGTATGGGTAATGC	1775						
Qy	1504	GGATCATGAAGATGGCACAC	1563						
Db	1776	GGATCATGAAGATGGCACAC	1835						
Qy	1564	TTGAGTCTTCTTGGTTCACAG	1623						
Db	1836	TTGAGTCTTCTTGGTTCACAG	1895						
Qy	1624	TAGTTATGTCTCTCGTGAAG	1683						
Db	1896	TGCTCTATGTTCTCTGTAAG	1955						
Qy	1684	TGAATGCCCTTGTTCGTGCTC	1743						
Db	1956	TGAATGCTCTGTTCGTGCTC	2015						
Qy	1744	ATTGTGATCACTACATCAACA	1803						
Db	2016	ATTGTGATCACTACATCAACA	2075						
Qy	1804	ATCCAACTAGCTCCGCAAGT	1863						
Db	2076	ACCTTACCTTAGGAAGAGTGT	2135						
Qy	1864	ATAGGAATGATCGATATGAAA	1923						
Db	2136	ACAGGAATGATCGATATGAAA	2195						
Qy	1924	TTGACGGCATTTCAAGACCAAGT	1983						
Db	2196	TTGATGGCATCCCAAGGACCAAGT	2255						

QY 156 GACGCTTTCACGGCCTGCGAGCTGTCGCCGCTTCCGGGTCTCGGCCCTGCTACAGAC 215
DB 334 GACGCTTTCACGGCCTGCGAGCTGTCGCCGCTTCCGGGTCTCGGCCCTGCTACAGAC 393
QY 216 GAGCCAAAGGAGGACCCAGCGCTGCTCCAGTGCAGAGACCAAGTACAACGCCACAGA 275
DB 394 GAGCCAAAGGAGGACCCAGCGCTGNNNNAGTGCNNNNCAAGTACAACGCCACAG 453
QY 276 GGGAGCCAGCGATCCCGGGGAGGAGGCGACGACACTGATGCCGATGATGGTAGTAC 335
DB 454 GGGAGTCCAGCGATCCCGAGGAGGAGGAGGACGATGCTGATGCCGATGATGCTAGCGAC 513
QY 336 TTCACTACCTGATCTGGCACTGAGGACGAGGAGGAGGATGCTGACAGATGCGC 395
DB 514 TTCACTACCTGATCTGGCACTGAGGACGAGGAGGAGGATGCTGACAGATGCGC 573
QY 396 AGCTGGCGCATGAACACCGGGGAGTGGCAATGTTGGCCACCACCAAGTATGACAGTGGC 455
DB 574 AGCTGGCGCATGAATGCTNNNNCAGCGGATGTTGGCCGCCCAAGTATGACAGTGGT 633
QY 456 GAGATCGGCCCTCCAAAGTATGACAGTGGAGAGATCCCTAGGGATAGCTCCCTTCAGTC 515
DB 634 GAGATCGGCCCTCCAAAGTATGACAGTGGAGATCCCTAGGGATAGCTCCCTTCAGTC 693
QY 516 ACCAACGCCAGATGTCAGGAGAAATCCCTGGAGCTTCCGCTGATCATCATGATGTC 575
DB 694 NNNNNNNNNATTTCCGGGAGAAATCCCTGGTCTTCCCTGACCATATGATGTC 753
QY 576 CCTAGGGGAAATCAGCAGAGCTGCTCCGTTTCCCTATGTAATCATACCAAAATCCG 635
DB 754 CCTAGTGGGAAATCAGCAGAGCTGCTCCGTTTCCCTATGTAATCATACCAAAATCCG 813
QY 636 TCAAGGAGTCTCCGGCAGTATGGGAATGTTGCCGGAAGAGAGATGATGGGTGG 695
DB 814 TCGAGGGAATCTCTGGTAGGTGGGAATGTTGCCGGAAGAGAGATGATGGGTGG 873
QY 696 AAATGAAGCAGGACAGGAGTCCGATCCCATGACTAATGGGACAGATGCTCCCTCT 755
DB 874 AAATGAAGCAGGACAGGAGTCCGATCCCATGACTAATGGGACAGATGCTCCCTCT 933
QY 756 GAAAGTCCGGCAGCTACTGACATCGATCTACTGAATCAACATGGAAGAGCGCTTA 815
DB 934 GAGGCGCGGCTGTTGGTGATTTGATGCACTCACTGATACACATGGAAGATGCGCTTA 993
QY 816 CTGAATGATGAATCTCGCAGCCTCTATCTAGAAAGTCCCATGCTTCCTCCAAATA 875
DB 994 TTAACGATGAATCTCGCAGCCTCTATCTAGAAAGTCCCATGCTTCCTCCAGGATA 1053
QY 876 AATCCCTACAGATGGTCTGCTGCGTGTGTTGTTGTTAAGCATCTTCCCTGCATAC 935
DB 1054 AATCCATACAGATGGTCTGCTGCTAGGATGATGTTCTAAGCATCTTCTTGCATAC 1113
QY 936 CGTCTCAAAATCTGTGCGTAATGCATACCCTGCTGGCTTTTATCTGTTATATGTAG 995
DB 1114 CGGATCAAAATCTGTGCGTAATGCATACCCTGCTGGCTTTTATCTGTTATATGTAG 1173
QY 996 ATTTGGTTGCTTTTCTCGATACCTGATCAGTCCGAGTGGTTCCCATCAACCGG 1055
DB 1174 ATCTGGTTGCTTTTCTCGATACCTGATCAGTCCGAGTGGTTCCCATCAACCGG 1233
QY 1056 GAGACCTACCTTGTAGACTGGCTTTAAGGTATGACCGAAGGTGAACGCTCTCAGTTG 1115
DB 1234 GAGACTTACCTTGTAGACTGGCTTTAAGGTATGACCGAAGGTGAACGCTCTCAGTTG 1293
QY 1116 GCTGCTGTGACATATTTGTCAGTACAGTGCACCCCTTGAAGGAGCCACCTATGCTCACT 1175
DB 1294 GCTGCTGTGACANNNNGTCAGTACTGTCAGCCCAATGAAGGAGCCCTCTTGTCACT 1353
QY 1176 GCCAACACTGTGCTATCCATCTTGTGCTGATTATCCCGTGGACAAAGGCTCTTGTCTAT 1235
DB 1354 GCCAATACCGTGTCTATCCATCTCGCTGTGGACTATCTCTGTGATAAGGCTCTGTCTAT 1413
QY 1236 GTATCTGATCAGGAGCTTCAATGCTGACTTTTGACGCAATGGCTGAGACTTTCAGAGTTT 1295

DB 1414 GTATCTGATGATGAGCTGCTA 1473
QY 1296 GCTAGGAAATGGGTACCATTTG 1355
DB 1474 GCTAGAAATGGGTGCCATTTG 1533
QY 1356 TACTTTTCCAGAAAATTTGATTT 1415
DB 1534 TACTTCTCCAGAAAATTTGATTT 1593
QY 1416 CGCCGGGCCATGAGAGAGATTT 1475
DB 1594 CGCCGGNNNNNNNNNNNNNNNN 1653
QY 1476 AAGCATTTGAAGTCCCGGAGG 1535
DB 1654 AAGCACNNNNNGTCCCTGAGG 1713
QY 1536 AACATACCAGGGATCATCTG 1595
DB 1714 AACAAATACCAGGGACCATCTG 1773
QY 1596 GATACTCAGGCTAATGAGCTCC 1655
DB 1774 GATACTCAGGCTAATGAGCTAC 1833
QY 1656 TTCAGCAGCACACAAGAGGCTG 1715
DB 1834 TTCAGCATCACAGAAGCTG 1893
QY 1716 ACTAATGGCAATACATGTTGA 1775
DB 1894 ACCAATGGCAATACATGTTGA 1953
QY 1776 GTCGAGAACTATGCTCTCC 1835
DB 1954 CTCAGGGAAGCTATGCTTCC 2013
QY 1836 CAGTTCCCAAAAGGTTTGATG 1895
DB 2014 CAGTTTCCCAAGGTTTCGATG 2073
QY 1896 GTCTTTTGTATTAATTAATTTGA 1955
DB 2074 GTGNNNNCATATTAATTTGA 2133
QY 1956 ACTGGTTGTCTTTTCAACAGAA 2015
DB 2134 ACTGGCTGTGTTTCAACAGAA 2193
QY 2016 AAGCCAGTTTCTTTGGCATCAC 2075
DB 2194 AAGGGTGTCTTGTCTATCAC 2252
QY 2076 AAGAGCTCAGATGAAGAAAATG 2135
DB 2253 -GGGCTCAGACAGNNNNNGN 2310
QY 2136 CTCGAAGACATAGAGGAGGTTG 2195
DB 2311 NNNNNNNNTAGAGGAGGAGT 2370
QY 2196 ATGCTCAAAATGAGCTTAGAGAA 2255
DB 2371 ATGCTCAAAATGAGCTTAGAGAA 2430
QY 2256 CTGATGAATATGGTGGTTCG 2315
DB 2431 CTGATGAATATNNNNNNNNNNNN 2490
QY 2316 ATCCATGTCTAAAGTTGTGGCTA 2375

Db	2491	ATCCATGTCATAAAGTTGTGGCTACGAGACAAGATTGAATGGGAATCGAGATTGGGTGG	2550
Qy	2376	ATCTATGGATCTGTCACAGAGATATTCTTACTGGATTCAAGATGTCACGCAAGAGGCTGG	2435
Db	2551	ATCTATGGTCTGTGACGGAAGATATTCTCACTGGGTTCAAGATGTCACGACGAGGCTGG	2610
Qy	2436	CGTTCACTATTCGATGCCCCAAGCGGCCAGCTTTCAAGGAGATCGCCCCCATCAATCTT	2495
Db	2611	CGGTGCGATCTACTGCATGCCATAAGCGGCCGCCCTCAAGGAGTCGGCTCCCATCAATCTC	2670
Qy	2496	TCAGATCGCTGAACCAAGTCCTCGGTGGGCTCTCGGTTCTGTGAAATCTTTTTCAGC	2555
Db	2671	TCAGACCGCTGAAACCAAGGTGCTCCGGTGGGCTCTCGGTTCAAGTGGAAATCTTTTTCAGC	2730
Qy	2556	CGGATTTGCCCTTATGGTATGGCTACGAGGGCGGCTCAAGTTTCCTGGAGAGATTCGCT	2615
Db	2731	CGGATTTGNNNNNTATGGTACGGGTACGGAGCGCCTGAAGTTCTTGGAGAGATTCGCC	2790
Qy	2616	TACATCAACACCAATTTTACCCACTAACCTCTCTCCCGCTTCTTAGTCTATTGTATATTG	2675
Db	2791	TACATCAACACCACTATCCCGCTCAGGTCCCTCCGCTCCCTCAATTACTGTATCCTG	2850
Qy	2676	CCTGCTATCTCTGCTCACTGGAAAGTTCAATCATGCCAGAGATAGCAACTTGGCCAGT	2735
Db	2851	CCTGCCATCTGCCCTGCTCAGGGGAAGTTCAATCATCCAGAGATCAGCAACTTCGCTAGT	2910
Qy	2736	ATCTGGTTCAPTGGCGCTTCTCTTCAATTTTCCGCACTGGTATCCCTTGAGATCAGGTGG	2795
Db	2911	ATCTGGTTCAPTGGCGCTTCTCTTCAATTTTCCGCACTGGTATCCCTTGAGATCAGGTGG	2970
Qy	2796	AGTGGTGTGGCATTCGAGAGTGTGGAGGAATGAACAGATTCTGGGTCATTGGAGGTATC	2855
Db	2971	AGCGCGTGGGCATCGACGAGTGTGGAGGAACAGCAGTTCCTGGTCACTCGGAGGCATC	3030
Qy	2856	TCTGCATCTGTTTGGCGTCTTCAGGGCTTCTGAAGTGCTTTCGGGTATCGACACC	2915
Db	3031	TCCGCCCACTCTTTCGGCGTCTTCAGGGGCTCCTCAAGGTGCTTGGCGGCATCGACACC	3090
Qy	2916	AACTTCACTGTCACCTCAAAGGCTAATGACGAAGCGCACTTGTGTGAGCTCTACATG	2975
Db	3091	AACTTCACTGTCACCTCAAAGGCTCGGATGAACAGCGCACTTTCGGGAGCTGTACATG	3150
Qy	2976	TTCAAGTGGAGCGCTTCTCACTCCCTCCGACGACCAATTTTGATCATTAACATGGTTGGT	3035
Db	3151	TTCAAGTGGAGCGACATCTGATCCCGCCACCACCATCCTGTGATCATCAACTGGTCGGC	3210
Qy	3036	GTCGTTGCTGCGACCTCTTACGCCATCAACAGTGGTTACCAATCATGGGGGCCCTCTTT	3095
Db	3211	GTTGTTGCGCGCATCTCTTACGCCATCAACAGCGGGTACCAGTCTGTGGGGTCCGCTCTC	3270
Qy	3096	GGGAAGCTCTTCTTGGCTTCTCGGTGATTGTTCACTTATACCACTTCCCTCAAGGGTCTT	3155
Db	3271	GGCAAGCTCTTCTTGGCTTCTCGGTGATTGTTCACTTATACCACTTCCCTCAAGGGTCTC	3330
Qy	3156	ATGGGAGGCAAAAACCGACACAGCATTTGTCATGCTGGGTGTCCTCTCCCTCGCTTCT	3215
Db	3331	ATGGGTGGGAGAACCGACCCCGACCATCGTGTGTTGCTGTGGGCGATCCTGCTGGCGTCG	3390
Qy	3216	ATCTTCTCTGCTGTGGGTCTGTTGATCCATTCACATCCCGTCTCGCTGGGCCCAAT	3275
Db	3391	ATCTTCTCTGCTGTGGGTCTGATCCGTTTCCATCCCAACCCGCTCACTTGGCCCGGAT	3450
Qy	3276	ATCCAAACCTGTGGCATCAACTGCTAGGAAAGTGGGA-GTTTGTAGAGACAGAAA	3329
Db	3451	ACTCGAAGGTGTGCATCAACTGTAGGAGGTGGAAGTTTGTAGAAACAGAGA	3505
RESULT	3		
LOCUS	AV110079	3696 bp	linear
DEFINITION	Zea mays CL1164_1 mRNA sequence.		
ACCESSION	AV110079		
VERSION	AV110079.1	GI:21214162	
			HTC 25-MAY-2002

KEYWORDS	HTC.
SOURCE	zea mays.
ORGANISM	zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	Hayney,C.F., Dolan,M.J., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL	Unpublished (2002)
REFERENCE	2 (bases 1 to 3696)
AUTHORS	Coe,E.C.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
FEATURES	Location/Qualifiers source 1..3696 /organism="Zea mays" /db_xref="MaizeDB:630055" /db_xref="taxon:4577" /clones="CU1184_1" /Clone_lib="Maize Mapping Project/DuPont Cornsensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"
BASE COUNT	831 a 766 c 880 g 945 t 274 others
ORIGIN	
Query Match	60.6%; Score 2199; DB 11; Length 3696;
Best Local Similarity	78.6%; Pred. No. 0;
Matches 2569; Conservative	0; Mismatches 686; Indels 12; Gaps 4;
Qy 38	GTTTGATCGCTGTCGGCCGCATGAGCGGACGCCGAGCCCTGAAGTCCGGGAGCA 97
Dd	
Db	232 GGAGGCTCGATTGTGGCCCATGAGCGGGC---GACCACCAGAAATTCGGGAAGCA 288
Qy 98	CGGGCGGGGACATGTGCCAGATCTGCGCGACGGCCTGGGCACCACTGTTGGACGGCA 157
Dd	
Db	289 TGTGCGGGCAGGTGTGCCAGATCTGCGCGACGGCGTGGCGGACCGCGCGGACGGCA 348
Qy 158	CGTCTTACCGCTCGACAGCTGTCGGCTTCCGGTCTCGCGCTCTGCTACGAGCACGA 217
Dd	
Db	349 CTTCTTACCGCCTCGCAGCTCTCGGCTTCCCGTGTGCGCGCATGCTACGAGTAGCA 408
Qy 218	CGCAAGGAGGGACACCGAGCCTGCCTCCACTGCAAGACCAAGTACAAGCCACACAGG 277
Dd	
Db	409 CGCAAGGACGGCACCCAGGCGTCCCGCAGTGCAGACTAAGTACAAGCCCAAGG 468
Qy 278	GAGCCCACGCTATCCGCGGGAGGAGCGACACACTGATCCCGATGTAGTGTGACTT 337
Dd	
Db	469 GAGCCCACCTACACGGTGAGGAAATAGAGGATGTGGATGCTGACGATGTGAGTGACTA 528
Qy 338	CAACTACCTGCATCTGGCACTGAGGACCAAGAGAAGATTCGTGACAGGATGGCGAG 397
Dd	
Db	529 CAACTACCAACACTCTGGCAACACAGGATCAGAAGCAAAGATTCGTGAGAATGCTCAC 588
Qy 398	CTGGCGATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACATGGCGA 457
Dd	
Db	589 TTGGCGCAACTACCTGGCAGTG---ATATTGGCTGGCTAAGTATGACAGCGGTGA 645
Qy 458	GATCGGCTCTCCAAGTATGACGTGGAGAGATCCCTAGGGGATACGCTCCCTTCAGTCAC 517
Dd	
Db	646 AATTGGGCATGGGAAGTATGACGTGGTGAGATCCCTCGTGATATATCCGCTCACTAAC 705
Qy 518	CAACAGCAGATGTACAGAGAAAATCCCTGGAGCTTCGCGCTGATCATCATGATGTCCCC 577
Dd	
Db	706 TCATAGCAGATCTCAGAGAGATCTCTGGAGCTTCGCCCTGATCATGATGCTCTCTGT 765

QY 578 TACGGGAACATCAGCAGACGTGCTCCGTTTCCTATGTGAATCATTTACCAATCCGTC 637
D 579 TACGGGAACATCAGCAGACGTGCTCCGTTTCCTATGTGAATCATTTACCAATCCGTC 638
D 580 TACGGGAACATCAGCAGACGTGCTCCGTTTCCTATGTGAATCATTTACCAATCCGTC 639
QY 638 AAGGAGTTCCTCGGAGTATTTGGGAATGTTGCTGGAAGAGAGAGTGTGATGCTGGAA 697
D 639 AAGGAGTTCCTCGGAGTATTTGGGAATGTTGCTGGAAGAGAGAGTGTGATGCTGGAA 698
D 640 AAGGAGTTCCTCGGAGTATTTGGGAATGTTGCTGGAAGAGAGAGTGTGATGCTGGAA 699
QY 698 AATGAACAGCAGCAAGGTCGCTATCCCATGACTTAATGGGACAGCAATTCCTCCTCTGA 757
D 699 AATGAACAGCAGCAAGGTCGCTATCCCATGACTTAATGGGACAGCAATTCCTCCTCTGA 758
D 700 AATGAACAGCAGCAAGGTCGCTATCCCATGACTTAATGGGACAGCAATTCCTCCTCTGA 759
QY 758 AGTTCGGGACGCTACGATGACATCGATGCTACTGATATCAACATGGAAGACGCTTTACT 817
D 759 AGTTCGGGACGCTACGATGACATCGATGCTACTGATATCAACATGGAAGACGCTTTACT 818
D 760 AGTTCGGGACGCTACGATGACATCGATGCTACTGATATCAACATGGAAGACGCTTTACT 819
QY 818 GAATGATGAACATCGCAGCCTCTATCTAGAAAAGTCCCATGCTTCCTCCCAATATAA 877
D 819 GAATGATGAACATCGCAGCCTCTATCTAGAAAAGTCCCATGCTTCCTCCCAATATAA 878
D 820 GAATGATGAACATCGCAGCCTCTATCTAGAAAAGTCCCATGCTTCCTCCCAATATAA 879
QY 878 TCCTACAGAAATGCTCATTTCTGCGGTTGTTGTTCTAAGCATCTTCCTGCACTACCG 937
D 879 TCCTACAGAAATGCTCATTTCTGCGGTTGTTGTTCTAAGCATCTTCCTGCACTACCG 938
D 880 TCCTACAGAAATGCTCATTTCTGCGGTTGTTGTTCTAAGCATCTTCCTGCACTACCG 939
QY 938 TCCTACAAATTCCTGCGGTTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGAGAT 997
D 939 TCCTACAAATTCCTGCGGTTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGAGAT 998
D 940 TCCTACAAATTCCTGCGGTTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGAGAT 999
QY 998 TTGCTTTGCTTTTCTGCTGATGCTGCTACGTTTGGCTGTTCTATGATGATTTCTGCTACCG 1122
D 999 TTGCTTTGCTTTTCTGCTGATGCTGCTACGTTTGGCTGTTCTATGATGATTTCTGCTACCG 1123
D 1000 TTGCTTTGCTTTTCTGCTGATGCTGCTACGTTTGGCTGTTCTATGATGATTTCTGCTACCG 1124
QY 1003 GAATGATGAACATCGCAGCCTCTATCTAGAAAAGTCCCATGCTTCCTCCCAATATAA 1062
D 1004 GAATGATGAACATCGCAGCCTCTATCTAGAAAAGTCCCATGCTTCCTCCCAATATAA 1063
D 1005 GAATGATGAACATCGCAGCCTCTATCTAGAAAAGTCCCATGCTTCCTCCCAATATAA 1064
QY 1063 TCCTACAGAAATGCTCATTTCTGCGGTTGTTGTTCTAAGCATCTTCCTGCACTACCG 937
D 1064 TCCTACAGAAATGCTCATTTCTGCGGTTGTTGTTCTAAGCATCTTCCTGCACTACCG 938
D 1065 TCCTACAGAAATGCTCATTTCTGCGGTTGTTGTTCTAAGCATCTTCCTGCACTACCG 939
QY 938 TCCTACAAATTCCTGCGGTTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGAGAT 997
D 939 TCCTACAAATTCCTGCGGTTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGAGAT 998
D 940 TCCTACAAATTCCTGCGGTTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGAGAT 999
QY 1123 TATCACACATCCTGTGAACATGCATACCCACTGTGGCTTTTATCTGCTCATATGTGAGAT 1182
D 1124 TATCACACATCCTGTGAACATGCATACCCACTGTGGCTTTTATCTGCTCATATGTGAGAT 1183
D 1125 TATCACACATCCTGTGAACATGCATACCCACTGTGGCTTTTATCTGCTCATATGTGAGAT 1184
QY 998 TTGCTTTGCTTTTCTGCTGATGCTGCTACGTTTGGCTGTTCTATGATGATTTCTGCTACCG 1122
D 999 TTGCTTTGCTTTTCTGCTGATGCTGCTACGTTTGGCTGTTCTATGATGATTTCTGCTACCG 1123
D 1000 TTGCTTTGCTTTTCTGCTGATGCTGCTACGTTTGGCTGTTCTATGATGATTTCTGCTACCG 1124
QY 1058 GACCTACCTTGATAGACTGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGC 1117
D 1059 GACCTACCTTGATAGACTGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGC 1118
D 1060 GACCTACCTTGATAGACTGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGC 1119
QY 1243 AACATACCTTGATAGACTGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGC 1302
D 1244 AACATACCTTGATAGACTGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGC 1303
D 1245 AACATACCTTGATAGACTGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGC 1304
QY 1118 TGTGTTGACATATTTGTCAGTACAGTTCGACCCCTTGAAGGAGCCACCTATGCTCACTGC 1177
D 1119 TGTGTTGACATATTTGTCAGTACAGTTCGACCCCTTGAAGGAGCCACCTATGCTCACTGC 1178
D 1120 TGTGTTGACATATTTGTCAGTACAGTTCGACCCCTTGAAGGAGCCACCTATGCTCACTGC 1179
QY 1303 TCCTGTTGATANNNGTCAGTACTGTGGATCCCAATGAAGGAGCCCTCCTGCTCACTGC 1362
D 1304 TCCTGTTGATANNNGTCAGTACTGTGGATCCCAATGAAGGAGCCCTCCTGCTCACTGC 1363
D 1305 TCCTGTTGATANNNGTCAGTACTGTGGATCCCAATGAAGGAGCCCTCCTGCTCACTGC 1364
QY 1178 CAACACTGTGCTATCCATCTCTGCTGTTGATATCCCGTGGACAAAGTCTCTGCTATGT 1237
D 1179 CAACACTGTGCTATCCATCTCTGCTGTTGATATCCCGTGGACAAAGTCTCTGCTATGT 1238
D 1180 CAACACTGTGCTATCCATCTCTGCTGTTGATATCCCGTGGACAAAGTCTCTGCTATGT 1239
QY 1363 AAATACCTGCTTTCCATCTCTGCTGCGATATTCGCGTTGCAAGGATATCTGCTATGT 1422
D 1364 AAATACCTGCTTTCCATCTCTGCTGCGATATTCGCGTTGCAAGGATATCTGCTATGT 1423
D 1365 AAATACCTGCTTTCCATCTCTGCTGCGATATTCGCGTTGCAAGGATATCTGCTATGT 1424
QY 1238 ATCTGATGACGAGCTTCAATGCTGACTTTTGACGATTTGGCTGAGACTTCAGAGTTGC 1297
D 1239 ATCTGATGACGAGCTTCAATGCTGACTTTTGACGATTTGGCTGAGACTTCAGAGTTGC 1298
D 1240 ATCTGATGACGAGCTTCAATGCTGACTTTTGACGATTTGGCTGAGACTTCAGAGTTGC 1299
QY 1298 TAGGAAATGGGTACCATTTGTTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTTA 1357
D 1299 TAGGAAATGGGTACCATTTGTTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTTA 1358
D 1300 TAGGAAATGGGTACCATTTGTTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTTA 1359
QY 1483 TAGAAAATGGGTTCCGTTCTGTGAAGATACAAATAGAGCCCTANGGCCCGGAAATGGTA 1542
D 1484 TAGAAAATGGGTTCCGTTCTGTGAAGATACAAATAGAGCCCTANGGCCCGGAAATGGTA 1543
D 1485 TAGAAAATGGGTTCCGTTCTGTGAAGATACAAATAGAGCCCTANGGCCCGGAAATGGTA 1544
QY 1358 CTTTTCGACAGAAATTTGATTTACCTGAAGCAAAAGTCCAGCTTCATTTGTTAAAGACCG 1417
D 1359 CTTTTCGACAGAAATTTGATTTACCTGAAGCAAAAGTCCAGCTTCATTTGTTAAAGACCG 1418
D 1360 CTTTTCGACAGAAATTTGATTTACCTGAAGCAAAAGTCCAGCTTCATTTGTTAAAGACCG 1419
QY 1543 CTTTGTCTAGAAAATTTGATTTACCTGAAGCAAAAGTTCAAACCTCATTTTGTGAAGAACG 1602
D 1544 CTTTGTCTAGAAAATTTGATTTACCTGAAGCAAAAGTTCAAACCTCATTTTGTGAAGAACG 1603
D 1545 CTTTGTCTAGAAAATTTGATTTACCTGAAGCAAAAGTTCAAACCTCATTTTGTGAAGAACG 1604
QY 1418 CCGGGCCATCAAGAGAGAAATATCAAGAAATTAATAATCAGGATAAATGCCCTAGTTCTAA 1477
D 1419 CCGGGCCATCAAGAGAGAAATATCAAGAAATTAATAATCAGGATAAATGCCCTAGTTCTAA 1478
D 1420 CCGGGCCATCAAGAGAGAAATATCAAGAAATTAATAATCAGGATAAATGCCCTAGTTCTAA 1479
QY 1603 CCGGGNN 1562
D 1604 CCGGGNN 1563
D 1605 CCGGGNN 1564
QY 1478 GGCATTTGAAGTCCCGAGGAGGATGGATGATGCAAGATGGCACACCAAGTGGCCAGGAA 1537
D 1479 GGCATTTGAAGTCCCGAGGAGGATGGATGATGCAAGATGGCACACCAAGTGGCCAGGAA 1538
D 1480 GGCATTTGAAGTCCCGAGGAGGATGGATGATGCAAGATGGCACACCAAGTGGCCAGGAA 1539
QY 1663 GGCACNN 1722
D 1664 GGCACNN 1723
D 1665 GGCACNN 1724
QY 1538 CAATACAGGAGTACCTCCGGAATGATTCAGGTTTTCCTTGGTCAGAGTGGGCTTGA 1597
D 1539 CAATACAGGAGTACCTCCGGAATGATTCAGGTTTTCCTTGGTCAGAGTGGGCTTGA 1598
D 1540 CAATACAGGAGTACCTCCGGAATGATTCAGGTTTTCCTTGGTCAGAGTGGGCTTGA 1599
QY 1723 CAATACAGGAGTACCTCCGGAATGATTCAGGTTTTCCTGAGGTTGACAGTGGGCTTGA 1782
D 1724 CAATACAGGAGTACCTCCGGAATGATTCAGGTTTTCCTGAGGTTGACAGTGGGCTTGA 1783
D 1725 CAATACAGGAGTACCTCCGGAATGATTCAGGTTTTCCTGAGGTTGACAGTGGGCTTGA 1784
QY 1598 TACTAGGGTAAATGAGCTCCCGCTTTAGTTTATGTCGTTGAAAAGGCTCCTGGGT 1657
D 1599 TACTAGGGTAAATGAGCTCCCGCTTTAGTTTATGTCGTTGAAAAGGCTCCTGGGT 1658
D 1600 TACTAGGGTAAATGAGCTCCCGCTTTAGTTTATGTCGTTGAAAAGGCTCCTGGGT 1659
QY 1783 CGTTGAGGCAATGAACTTCCCTGTTGGTTTATGTCGTTGNNNNNNNNNNNNNNNNNN 1842
D 1784 CGTTGAGGCAATGAACTTCCCTGTTGGTTTATGTCGTTGNNNNNNNNNNNNNNNNNN 1843
D 1785 CGTTGAGGCAATGAACTTCCCTGTTGGTTTATGTCGTTGNNNNNNNNNNNNNNNNNN 1844

QY 1658 CCAGCACCAACAAGAGGCTGGT 1717
D 1659 CCAGCACCAACAAGAGGCTGGT 1718
D 1660 CCAGCACCAACAAGAGGCTGGT 1719
QY 1718 TAATGGCAATACATGTTGAAG 1777
D 1719 TAATGGCAATACATGTTGAAG 1778
D 1720 TAATGGCAATACATGTTGAAG 1779
QY 1903 TAATGGCAATACATGTTGAAG 1962
D 1904 TAATGGCAATACATGTTGAAG 1963
D 1905 TAATGGCAATACATGTTGAAG 1964
QY 1778 CCAGAGAGCTATGTTGCTTCT 1837
D 1779 CCAGAGAGCTATGTTGCTTCT 1838
D 1780 CCAGAGAGCTATGTTGCTTCT 1839
QY 1963 TCAGAGAGCTATGTTGCTTCT 2022
D 1964 TCAGAGAGCTATGTTGCTTCT 2023
D 1965 TCAGAGAGCTATGTTGCTTCT 2024
QY 1838 GTTCCCAAAAGGTTTTCATG 1897
D 1839 GTTCCCAAAAGGTTTTCATG 1898
D 1840 GTTCCCAAAAGGTTTTCATG 1899
QY 2023 ATTTCCCTCAGAGGTTTTCATG 2082
D 2024 ATTTCCCTCAGAGGTTTTCATG 2083
D 2025 ATTTCCCTCAGAGGTTTTCATG 2084
QY 1898 CTTTTTTGATATTAACATTGAG 1957
D 1899 CTTTTTTGATATTAACATTGAG 1958
D 1900 CTTTTTTGATATTAACATTGAG 1959
QY 2083 GNNNNNGATATTAACATTGAG 2142
D 2084 GNNNNNGATATTAACATTGAG 2143
D 2085 GNNNNNGATATTAACATTGAG 2144
QY 1958 TGGTTGTTTTCACACAGAAAG 2017
D 1959 TGGTTGTTTTCACACAGAAAG 2018
D 1960 TGGTTGTTTTCACACAGAAAG 2019
QY 2143 TGGTTGTTGTTTTCACACAGAA 2202
D 2144 TGGTTGTTGTTTTCACACAGAA 2203
D 2145 TGGTTGTTGTTTTCACACAGAA 2204
QY 2018 GCCAGGTTTTCATGCACTACT 2077
D 2019 GCCAGGTTTTCATGCACTACT 2078
D 2020 GCCAGGTTTTCATGCACTACT 2079
QY 2203 GCCAGGTTTTCATGCACTACT 2259
D 2204 GCCAGGTTTTCATGCACTACT 2260
D 2205 GCCAGGTTTTCATGCACTACT 2261
QY 2078 GAGCTCAGATTAAGAAAGTCC 2137
D 2079 GAGCTCAGATTAAGAAAGTCC 2138
D 2080 GAGCTCAGATTAAGAAAGTCC 2139
QY 2260 GAGCTCAGATTAAGAAAGTCC 2319
D 2261 GAGCTCAGATTAAGAAAGTCC 2320
D 2262 GAGCTCAGATTAAGAAAGTCC 2321
QY 2138 CGAAGACATAGAGAGGTTGTT 2197
D 2139 CGAAGACATAGAGAGGTTGTT 2198
D 2140 CGAAGACATAGAGAGGTTGTT 2199
QY 2320 CGAAGACATAGAGAGGTTGTT 2379
D 2321 CGAAGACATAGAGAGGTTGTT 2380
D 2322 CGAAGACATAGAGAGGTTGTT 2381
QY 2198 GTCTCAATGAGCTTAGAGAG 2257
D 2199 GTCTCAATGAGCTTAGAGAG 2258
D 2200 GTCTCAATGAGCTTAGAGAG 2259
QY 2380 GTCTCAATGAGCTTAGAGAG 2439
D 2381 GTCTCAATGAGCTTAGAGAG 2440
D 2382 GTCTCAATGAGCTTAGAGAG 2441
QY 2258 GATGGATATGTTGGTTCTCT 2317
D 2259 GATGGATATGTTGGTTCTCT 2318
D 2260 GATGGATATGTTGGTTCTCT 2319
QY 2440 GATGGATATGTTGGTTCTCT 2499
D 2441 GATGGATATGTTGGTTCTCT 2500
D 2442 GATGGATATGTTGGTTCTCT 2501
QY 2318 CCATGTCATAGTTTGTGGCTAT 2377
D 2319 CCATGTCATAGTTTGTGGCTAT 2378
D 2320 CCATGTCATAGTTTGTGGCTAT 2379
QY 2500 TCATGTCATAGCTTGGCTAT 2559
D 2501 TCATGTCATAGCTTGGCTAT 2560
D 2502 TCATGTCATAGCTTGGCTAT 2561
QY 2378 CTATGGATCTGTCACAGAGAT 2437
D 2379 CTATGGATCTGTCACAGAGAT 2438
D 2380 CTATGGATCTGTCACAGAGAT 2439
QY 2560 CTATGGTTCTGTTACAGAGAG 2619
D 2561 CTATGGTTCTGTTACAGAGAG 2620
D 2562 CTATGGTTCTGTTACAGAGAG 2621
QY 2438 TTGAGTCTATGTCATGCCAAG 2497
D 2439 TTGAGTCTATGTCATGCCAAG 2498
D 2440 TTGAGTCTATGTCATGCCAAG 2499
QY 2620 NNNNNNNNNNNNNNNNNNNNNNN 2679
D 2621 NNNNNNNNNNNNNNNNNNNNNNN 2680
D 2622 NNNNNNNNNNNNNNNNNNNNNNN 2681
QY 2498 AGATCGTCTGAACCAAGTCTG 2557
D 2499 AGATCGTCTGAACCAAGTCTG 2558
D 2500 AGATCGTCTGAACCAAGTCTG 2559
QY 2680 GGATCGTTGAATCAAGTGGTT 2739
D 2681 GGATCGTTGAATCAAGTGGTT 2740
D 2682 GGATCGTTGAATCAAGTGGTT 2741
QY 2558 GCATGCCCCCTTATGTTATGGG 2617
D 2559 GCATGCCCCCTTATGTTATGGG 2618
D 2560 GCATGCCCCCTTATGTTATGGG 2619
QY 2740 GCATGCCCCCTTATGTTATGGG 2799
D 2741 GCATGCCCCCTTATGTTATGGG 2800
D 2742 GCATGCCCCCTTATGTTATGGG 2801
QY 2618 CATCAACACCACTTTTACCC 2677
D 2619 CATCAACACCACTTTTACCC 2678
D 2620 CATCAACACCACTTTTACCC 2679
QY 2800 TATCAACACCACTTTTATCC 2859
D 2801 TATCAACACCACTTTTATCC 2860
D 2802 TATCAACACCACTTTTATCC 2861
QY 2678 TGTATCTGCTGCTCACTGGA 2737
D 2679 TGTATCTGCTGCTCACTGGA 2738
D 2680 TGTATCTGCTGCTCACTGGA 2739
QY 2860 AGCAGTTTGTCTCTCACTGG 2919
D 2861 AGCAGTTTGTCTCTCACTGG 2920
D 2862 AGCAGTTTGTCTCTCACTGG 2921
QY 2738 CTGGTTCAATGCGCTCTTCT 2797
D 2739 CTGGTTCAATGCGCTCTTCT 2798
D 2740 CTGGTTCAATGCGCTCTTCT 2799

Db	2920	TTGGTTTATATCGCTCTTTATCTCAATCTTTGCCACTGGTATCCTTTGNNNNNNNNNN	2979
Qy	2798	TGCTGTGGCATTTGACGAGTGGTGGAGGAATGAACAGTTCCTGGGTCAATTTGGAGGTATCTC	2857
Db	2980	NNNNNNNNCATTGATGAATGCTGGAGGAACGAGCTTCCTGGGTCAATGGTGGTATTTCT	3039
Qy	2858	TGCACATCTGTTTGGCGCTTTTCAGGGCTTCCTGAAGTGGCTTGGCGGTATCGACACAA	2917
Db	3040	TGGCATTTTATTGGCGCTTCTCCAGGGTCTCCTGAAGTGGTCTGGTATCGACACGAG	3099
Qy	2918	CTTCACTGTCACTCAAAAGGCTTAATGACGAAGAGGCGACATTTGCTGAGCTCTACATGTT	2977
Db	3100	CTTCACTGTCACTCTAAGGCCACTGACAGAAGGTGATTTTGGCAGGCTCTACATGTT	3159
Qy	2978	CAAGTGGACGAGCGTTCTCATCCCTCCGACGACCATTTTGTATCATTAACATGGTTGGTGT	3037
Db	3160	CAAGTGGACGAGCGTTCTGATCCCAACCAACCACTATTTTGATCATCAACTGGTGGCGGT	3219
Qy	3038	CGTTGCTGGCACTCCTTAGCGCATCAACAGTGGTTACCAATCATGGGGCGGCTCTTTGG	3097
Db	3220	GGTGGCTGGCATTTCTACGCAATCAATAGCGGGTTACAGTCAATGGGGAACCTCTTTTCGG	3279
Qy	3098	GAAGCTCTTCTTTGGCTCTCTGGGTGATTTGTTCACATTAACCCATTCCTCAAGGCTCTTAT	3157
Db	3280	GAAGCTCTTCTTTGGTCTCTGGTGAATTTGCCACTGTACCCCTTCTCTCAAGGGCCAT	3339
Qy	3158	GGCAGGCAAAACCGACACACGACGATTTGTCATGCTGTGGGCTGTCTCTCTCGCTTCAT	3217
Db	3340	GGGAGAGCAGAAACCGACCGCCGACCATTTGTGTTCTCTGGGCTATCCTCTCTGGCTCGAT	3399
Qy	3218	CTTCTCCTTGTCTGGGTTCGTGTTGATCCATTCACTACCCGCTCTCGCTGGCCCAATAT	3277
Db	3400	CTTTCCCTTGATGTGGGTTCGTATCGATCCATTCACCAACCGGGTCACTGGTCTCTGATAT	3459
Qy	3278	CCAAAGCTGTGGCATCAACTGCTAGGA	3304
Db	3460	CGCGAAATGTGGCATCAACTGCTAGGA	3486

RESULT 4	LOCUS	DEFINITION	3898 bp	mRNA	linear	HTC 25-MAY-2002
AY110415	zeamays	CL1166.1 mRNA sequence.				

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.

BASE COUNT		932 a		841 c		991 g		1054 t		80 others		
ORIGIN												
contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"												
Query Match	37.9%	Score	1373.2	DB	11	Length	3898					
Best Local Similarity	65.7%	Pred.	No. 0									
Matches 2110	Conservative	0	Mismatches	1005	Indels	98	Gaps					
QY	53	CCGCGCATGACGGCGGACGCGGACGCCCTGAAGTCCGGGAGGCACGGGCGCGGGACGT	112									
Db	248	CCGCCAGCAGCGGCGCGCGCGCGGCTAAGCCACGAAGAGTGGGAATGGCAGGT	307									
QY	113	GTGCCAGATCTGGCGGAGGCGCTGGGCACACAGTTGAGCGGCAGCTTTCACCGCGTG	172									
Db	308	CTGCCAGATTTGGCGCGAGCTCTGTGGGGGTTTCAGCCACTGGTGATGCTTTGTGTCCTG	367									
QY	173	CGACGTCCTCCGCTTCGCGCTTCGCCCTCTGCTAGCAGCACGACGCGAAGGAGGCAC	232									
Db	368	CAATGAGTGTGCTTCCCTGCTGCGGCTTCTGTATGAGTACGCGCGAAGGAGGGAA	427									
QY	233	CCAGGCGCTGCCCTCCAGTGCACACCAAGTACAAGCGCCACAGAGGAGCCCGCGATCCG	292									
Db	428	CCAATGCTGCCCTCAGTGCAGACATAGATACAAGAGACAGAAAGTAGGCCCTCGAGTTCA	487									
QY	293	CGGGGAGGAGGCGGACGACACTGATGCCGATGATGTTAGTGACTTCAACTACCTCGATC	352									
Db	488	TGTTGATGAGGATGAGGAAGATGTTTGATGACTAGACAATGAATTCACATACAAGCAAGG	547									
QY	353	TGGCACTGAGGACACAGACAGAAATGCTGCACAGGATGCCAGCTGGCGGCATGAACAC	412									
Db	548	CAGTGGGAAGGGCCAGAGTGCAGCT-----GCAAGAGGAT	584									
QY	413	CGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCTCTCCAA	472									
Db	585	GAGCGTGATCTGTCTTCATCTGCTCGCCATGCCACACCATCGGATTCACGCCCTGACA	644									
QY	473	GTATGACAGTGGAGAGATCCCTAGGGGATAGTGCCCTTCAGTCAACACGCGAGATGTC	532									
Db	645	AGCGGTCA-----ACAGATATC	661									
QY	533	AGAGAAATCCCTGGAGCTTCGCTGATCATCATGATGTCCCTACGGGGGAACATCAG	592									
Db	662	TGAGAGATCCCTGATGATCCCTGACCGCTATCTATCCGAGTCCCAACATCGAGCTA	721									
QY	593	CAGACGTGCTCGGTTTCCCTATGTGAATCATTCACAAATCCGTCAAGGAGTT---CTC	649									
Db	722	TGTTGATCAAGCGTCCAGTCCTGTGAGGATTTGGACCCCTCGAAGGACTTGAATTC	781									
QY	650	CGGCAGTATGGGAATGTTGCCGTGGAAGAGAGAGATTGATGGCTGGAAAATGAAGCAGGA	709									
Db	782	CTATGGCTTAATAGTGTGACTGGAAGAAAGAGTTGAGAGCTGGAGGTTTAAACAGGA	841									
QY	710	CAAGGTGCGAATCCCATGACTAATGGGACAAGCATTTGCTCCCTCTGAAGTCGGGCAGC	769									
Db	842	CNNNNATGTTGCAAGTGACTAATAAATATCCAGAGGCTAGAGGAGGACATGAGGAGG	901									
QY	770	TACTGACATCGATGCTACTGTAATACAACATGAAGACGCTTTACTGAATGATGAAC	829									
Db	902	GACTGCTCAAAATGGA-----GAAGATATGCAAAATGTTGATGATGC	943									
QY	830	TGCCAGCCTCTATCTAGAAAAGTCCCATTTGCTTCTCCAAAATAAATCCCTACAGAT	889									
Db	944	ACGCTTACCTTTGAGCCGATCGTGCCAAATTTCTCAAAACGAGCTCAACCTTTACCGGT	1003									
QY	890	GGTCAATGTCGCGGTTGGTTGTCTAAGCATCTGCTGACCTACCGTCTCAACAATCC	949									
Db	1004	AGTGATCATCTCCGCTTATCATCTGCTGCTTCTTCCAGTATCGTATCAGTCATCC	1063									
QY	950	TGTCGCTAATGCAATACCACCTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTTCT	1009									

Db 1064 AGTCGCTGATGCTTATGGATTGTGGCTAGTATCTGTATCTCGAGGCTGGTTGCGCTT 1123
QY 1010 ATCCTGGATACTGATCAGTTCCTCCGAGTGGTTTCCAATCAACCGGGAGACCTACCTTGA 1069
Db 1124 GTCTTGGCTTCTAGATCAGTTCCTCCAAATGGTATCCAATCAACCGGGAGACATATCTCGA 1183
QY 1070 TAGACTGGCTTTAAGGTATGACCGAAGGTGACCGTCTCAAGTGTGGCTGCTGTGACAT 1129
Db 1184 CAGGCTTGCAATCAGGTATGATAGAGAGGAGAGCCATCACAGCTGGCTCCCATGTATGT 1243
QY 1130 ATTTGTCAGTACAGTCGACCCCTTGAAGGAGCCACCTATCGTCACCTGCCAACACTGTGCT 1189
Db 1244 CTTGCTGAGTACAGTGGATCCATTGAAGGAACCTCCACTGATCAGACGCAACACTGTGTTT 1303
QY 1190 ATCCATCTGCTGTTGATATCCCTGGACAGGCTCTCTGTATGCTATGCTATGATGACGG 1249
Db 1304 GTCCATCTTGTGCTGATACCTGTTGACAAGGTGCTGATGCTATGTTCTGATGAGG 1363
QY 1250 AGCTTCAATGCTGACTTTTGAACGATTTGGCTGAGACTTTCAGAGTGTGGCTAGGAAATGGT 1309
Db 1364 CTCAGCTATGCTGACTTTTGAAGTCTCTCTGAAACCGCCGAATTTGCTAGAAAGTGGT 1423
QY 1310 ACCATTTGTAAGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTTACHTTTTGGCCAGAA 1369
Db 1424 TCCTTTTGTGAAGGACACAATATGAACCAAGAGCTCCAGAAATTTTACTTTGCTCNNN 1483
QY 1370 AATTGATTACCTGAAGACAAAGTCCAGCCTTCATTTGTTAAAGACCGCGGCCATGAA 1429
Db 1484 NNTAGATTACCTGAAGGACAAAATTCACCTTCATTTGTTAAAGAAAGACGCGCAATGAA 1543
QY 1430 GAGAGAAATGAGAATTTAAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTTGAAGT 1489
Db 1544 GAGGAGTATGAAGATTCAGAAATGAAGTCAATGCCCTGTTTGGCCAAAGCACAGAAAGT 1603
QY 1490 CCCCAGGAGGATGATCATGATCAAGATGACACACCATGCCAGGAAACATACCCAGGA 1549
Db 1604 GCCTGAAGAGGGTGGACCATGCTGATGGAAGTCTTGGCCTGGGAATTAACCTTAGGGA 1663
QY 1550 TCATCTCGGAATGATTCAGGTTTTCCTTGTGTCACAGTGGTGGCTTCATCTAGGGTAA 1609
Db 1664 CCATCTGGCATGATTCAGGTGTTCTTGGGGCAGAGTGGTGGCTCGACACTGATGGA 1723
QY 1610 TGAGTCCCCGTTTATGTTATGTTCTCGTGAAAGCGTCTCGGGTTCAGCACCCACAA 1669
Db 1724 TGAGTTACCAAGCTTCTGCTATGCTCTCGTGAAAGAGACCAAGCTTTCAGCATCACA 1783
QY 1670 GAAGGCTGGTCCATGAATGCCCTTGTCTGTCACAGTGTCTCTACTAATGGACAATA 1729
Db 1784 GAAGGCTGGTGCAATGCAATGCGCTGATTCGTTGATCTGCTGCTGCAAAATGGTGCTTA 1843
QY 1730 CATGTTGAATCTTGATTTGATCAGTACATCAACACAGCAAGGCTCTCGGAGAAGCTAT 1789
Db 1844 TCTTCTCAATGTGATTTGGACCATTAATCAATAGCAGCAAAAGCTCTTAGAGAAGCAAT 1903
QY 1790 GTGTTTCCATTAATGATCCAAACCTAGTTCGCGAAGTCTGTTATGTCAGTTTCCCAACAA 1849
Db 1904 GTGCTTCATGATGATCCGCTCTAGGAAGGAAACCTGTTATGTACAAATTTCCACAGAG 1963
QY 1850 GTTGTGAGGATGATAGGAATGATCGATATGCAACAGCAAGCACTGCTTTTTTGTATAT 1909
Db 1964 ATTTGATGGCATTGACTTGCAGCATCGATATGCTAATCGGAACATAGTCTCTTTGATAT 2023
QY 1910 TAACTTGAGGGGCTTGACGGCATTTCAAGGACCAAGTTTATGTGGGAAGTGTGTTT 1969
Db 2024 CAACATGAAAGGCTGGATGGCATTCAGGGTCCAGTCTAGCTGGGAACAGGATGCTGTTT 2083
QY 1970 CAACAGAACAGCTATCTATGTTATGAGCCCCCAATTAAGCGGAAGAGCCAGGTTTCTT 2029
Db 2084 CAATAGGACGCTTGTATGGATACGA-TCCGTGTTTGTAGTGAAGCTGATCTGGAGCCTA 2142
QY 2030 GGCATCACTATGTGGGGGCAAGAGAAAGCAAGTCAAAAGAAAAGGAGCTCAGATAA 2089
Db 2143 ACATTTGTTTGAAGAGCTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATATGATA 2202

QY 2090 GAAAAAGTCGAACAAGCATGCT 2145
Db 2203 GTCANAGCCGTATATGAAGAG 2262
QY 2146 TAGAGGAGGTGTGGAAGGTG 2205
Db 2263 TCGAGGAGGTGTGGAAGGT- 2316
QY 2206 TGAGCTTAGAGAAGAGATTG 2265
Db 2317 GSAATTTGGAGAAGCGTTTG 2376
QY 2266 ATGCTGTGTTCCTCAAGTCT 2325
Db 2377 AAGGTGCATACCACTTCAAG 2436
QY 2326 TAAAGTTGGCTATGAGGACA 2385
Db 2437 TCAGTTTGGATATGAGGACA 2496
QY 2386 CTGTCACAGAAGATATTCTTA 2445
Db 2497 CAGTTACGGAGGATATTCTGA 2556
QY 2446 ATTGCATGCCCAAGCGCCAG 2505
Db 2557 ACTGCATGCCACCAAGACCTT 2616
QY 2506 TGAACCAAGTGTGCGGTGGG 2565
Db 2617 TTAATCAGAGTGTGCGTGGG 2676
QY 2566 CCTATGCTATGGCTACCGAG 2625
Db 2677 CTATCTGTATGTTTCAATG 2736
QY 2626 CCACCATTTCCCACTAACCT 2685
Db 2737 CCATTGTTTATCCAATCACAT 2796
QY 2686 GTCTGCTCAGTGGAAAGTTCA 2745
Db 2797 GTCTTCTTACCAATAAATTA 2856
QY 2746 TTGCGCTCTTCTTCAATTT 2805
Db 2857 TTCTTCTTTTGGCCTCCANN 2916
QY 2806 GCATTGACGAGTGTGAGGA 2865
Db 2917 GCATTGAAGATTTGGTGAGAA 2976
QY 2866 TGTGTCGCTTCTTTCAGGGTC 2925
Db 2977 TCTTCGAGTGTTCAGGGTC 3036
QY 2926 TCACCTCAAAGGCTTAATGACG 2985
Db 3037 TTACCTCAAAGGCAATCGATG 3096
QY 2986 CGAGCTTCTCATCCCTCCGA 3045
Db 3097 CCAGTTTGCATATTCCTCCGA 3156
QY 3046 GCACCTCTACGCCATCAACA 3105
Db 3157 GAATTTGATGATCCATTAACA 3216
QY 3106 TCTTTGCTTCTGGGTGATG 3165
Db 3217 TCTTCTGATCTGGGTGNTCC 3276

Db	1388	AAGAGCTTGGAGAAACGCTTTGGCCAGTCTCCAAATTTTATTGCAATCACCTTTATGACT	1447
Qy	2265	TATGGTGGTCTTCCTCAGTCCTCCACTCCAGAAATCTCTTTTGAAGAAGCTATCCATGTC	2324
Db	1448	CAAGGTGGCATACCCCTTCAACAACCAGGTTCCTTGCTTAAGGAAGCTATACATGTC	1507
Qy	2325	ATAAGTTGTGGCTATGAGGACAAGTCTGAATGGGAACTGAGATTGGTATATATGGA	2384
Db	1508	ATTAGTTGTGGATATCAGGATATAACAGAAATCGGGGAAAGAGATCGGATGATATATGCG	1567
Qy	2385	TCGTGCACAGAATATCTTACTGGATTCAGATGCACGAAGAGCTGCGGTTCAGTC	2444
Db	1568	TCGTGTACTGAAGATATTTAACTGGTTTCAAGATGCGATCGAAGAGTTGGATATCCATC	1627
Qy	2445	TATTGCATGCCAAGCGCCAGCTTTCAAGGATCTGCCCCATCAATCTTTTCAGATCGT	2504
Db	1628	TACTGCATGCCACTTCGGCTTGCTTCAAGGGTCTGCTCCGATTAAATCTTTCTGATCGT	1687
Qy	2505	CTGNACCAAGTCTGCGGTGGGCTCTCGGTTCTGTGAAATCTTTTCAGCCGGCATGTC	2564
Db	1688	CTCAACCAAGTGTACGCTGGGCTCTGGTTCAGTTGAAATTTCTACTTAGCAGACACTGT	1747
Qy	2565	CCCTTATGTTATGGCTACGAGGGGGGCCCTCAAGTTCCCTGGAGAGATTCGCTTACATCAAC	2624
Db	1748	CCATCTGGTATGTTTACAATGGAAGCTTAAGCTTCTGGAGAGACTGGCATACATCAAC	1807
Qy	2625	ACGACCAATTACCCATACCTCTCTCCGGCTTCCTAGTCTATATTGTAATTCGCTGCTATC	2684
Db	1808	ACCATTTGTTTATCCAAATTCATCTATCCACCTAGTAGGATACTGCGTCTTCCCTGCTATC	1867
Qy	2685	TGCTGCTCTACTGGAAGTTTCATATGCCAGAGATTAGCAACTTGGCCAGTATCTGCTTC	2744
Db	1868	TGTTTACTACCAACAATTTAATATCTCGCATTTAGCAATATATGCTGGGGCGTCTCTC	1927
Qy	2745	ATTGCGCTCTCTTTCAATTTTCGCCACTGTTATCTCTTGAGATGAGGTGGAGTGGTGT	2804
Db	1928	ATCTGCTTTTTCCTTCCATCTTCGCCACTGTTATTTTGGAGCTTCGATGGAGTGGTGT	1987
Qy	2805	GGATTACGAGTGTGGAGGAATGAACAGTTCTCGGTCAATGGAGCTATCTCTGCACAT	2864
Db	1988	GGCATGAGGATTTGGTGGAAATGAGCAGTTTGGTTCATTTGGTGGCCACTCTGCACAT	2047
Qy	2865	CTGTTGCGCTCTTTCAGGCTCTCTGAAAGTGTCTTCCGGTATCGCACACAACCTTCAC	2924
Db	2048	CTCTTTGCTGTGTCCAAGGTCTTTAAAGTGTCTAGCAGGGATCGACACAAACTTCACG	2107
Qy	2925	GTCACTCTAAAGCTAATGACGAAGAAGCGACTTTGCTGAGCTCTACATGCTCAAGTGG	2984
Db	2108	GTCAATCAAAAGCAACCGATGATGTGATTTGCTGAGCTGATGTGTTTCCANGTGG	2167
Qy	2985	ACGAGCTTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGGTTGGTCTGCTTGT	3044
Db	2168	ACAACTCTTGATCCGCCCCACCACTGTGCTGTGTGATTAACTGGTTGGTATAGTGGCT	2227
Qy	3045	GGCACTCTTACCCCATACAGTGGTTACCAATCATGGGGGCCCTCTTTTGGGAAGCTC	3104
Db	2228	GGAGTGTGCTATCATCAACAGTGGCTACCAATCATGGGTCCACTATTCGGGAAGCTG	2287
Qy	3105	TTCTTTGCCCTCTTGGGTGATTTGTTCACTTATACCAATTCCTCAAGGGTCTTATGGGAGG	3164
Db	2288	TTCTTTGCAATCTGGGTGATCTCCACCTCTACCTCTTCCCTCAAGGGTCTCATGGGAAG	2347
Qy	3165	CANAACGCACACCGAGATTTGTATCTGCTGGGCTGTCTCTCGCTTCTATCTCTCC	3224
Db	2348	CAGAACGCACACCGACCACTGTCAATTTGTTGGTCCGCTCTTCTTGTCTTCCATATTCG	2407
Qy	3225	TTGCTGTGGTTCGTTGATCCATTCACACCGT	3260
Db	2408	CTGCTGTGGGTGAAGATCGACCCCTTCATATCCCT	2443
RESULT 6			

Db 703 ACATCCCGCGGAGCAGCAGCGCTCGTCCGCTCTACATGGCGCGCGCGGCGCA 762
QY 566 CATGATGTCCTCCCTACGGGGAACATCAGCAGAGTGCCTCGTTCCCTATGTGAATCATTC 625
Db 763 AGAGGATCCACCCGCTCCCTCGCTGATTCACACCTCCAGTGCAACCGAGATCCATGG 822
QY 626 ACCAAATCCGTCGAAGGAGTTCCTCCGGCAGTATTTGGGAATGTTGCCGTGGAAGAGAGAGT 685
Db 823 ACCCGTCCAAGGA--TCTGGCGCCCTACGSAATGCGACGGTGGCTGGAAGAGAGAT 880
QY 686 TGATGGCTGGAATAAGACGAGGAAAGGGTGGGATTCGCCATGACTAATATGGGACAGCAT 745
Db 881 GGAGGGCTGGAAGCAGAGCAGGAGCGCTGCGAGATGTCAGGAGCGAGG----- 931
QY 746 TGCTCCCTCTGAAGGTCGGCAGCTACTGACATCGATGCATCTACTGAATACAACATGGA 805
Db 932 -----TGCGGTGATGGGATGGCGAGCATGC 958
QY 806 AGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTATCTAGAAAAGTCCCAATTGCTTC 865
Db 959 AGATCTGCCACTAATGGATGAAGCTAGCGAGCAATGTCAGAAAAGTCCCTATATCATC 1018
QY 866 CTCCAAATAATCCCTACAGAATGTCATGTTCTGCGGTGGTGGTTCCTAAGCATCTT 925
Db 1019 AAGCCGAATTAATCCCTACAGGATGATATCGTTATCCGGTTGGTGGTTTGGGTTCTT 1078
QY 926 CTTGCACTACCGTCTCACAAATCCTCTGGCTAATGCATACCCACCTGTGCTTTATCTGT 985
Db 1079 CTTCCACTACCGAGTATGCATCCGGGAAAGATGATTTGCATTTGGCTCATATCTGT 1138
QY 986 TATATGTGAGATTTGGTTTGTCTTATCTCGTGATCTGGATCTAGTTCCCGAAGTGGTTTCC 1045
Db 1139 AATCTGTGAATCTGGTTGGGATGTCCTGGATCTTGATCAGTCCCAAGTGGCTTCC 1198
QY 1046 AATCAACCGGGAGACCTACCTTGATAGACTGGCTTTAAGGTATGACGAGAGGTGAACC 1105
Db 1199 AATCGAGAGAGACTTACTTGCACGCTTGTCTACTAAGTTTGAAGAAGGTCAACC 1258
QY 1106 GTCTCAGTTGGCTGCTGTGACATATTTGTCAGTACAGTCGACCCCTTGAAGAGCCACC 1165
Db 1259 CTTCTAGTTGGCTTCCCAATCGACTTCTTGTCAGTACGGTTGATPCCCAAGAAAGCCTCC 1318
QY 1166 TATCGTCACTGCCAACACACTGTGCTATCCATCTTCTGTTGATATCCCGTGACAAAGT 1225
Db 1319 CTTGGTCACAGCGAACACTGCTCTTCCATCCTTCTGTTGATATCCCGTTGAGAGGT 1378
QY 1226 CTTCTGCTATGTATCTGATGACGAGCTTCAATGCTGACTTTTTCGCAATGGCTGAGAC 1285
Db 1379 CTTCTGCTATGTTCTGATGATGGTGTGCAATGCTTACGTTTGAAGCATTTGTCTGAAC 1438
QY 1286 TTCAGAGTTTGGCTAGGAATGGGTACCATTTGTTGAAGAGATGATGACATTTGAACCCAGAC 1345
Db 1439 ATCTGAATTTGCAAGAAATGGGTTCTCTTCAGCAAAAAGTTTAATATCAGCCTCGTGC 1498
QY 1346 TCCCGAGTTTACTTTTCCAGAAAATTGATTAACCTGAAAGACAAAAGTCAGCCTTCATT 1405
Db 1499 TCTGAGTGTACTTCCACAGAGAATAGACTACCTTGAAGACAGAGTTGCTGCTTCATT 1558
QY 1406 TGTAAAGACCGCGGCCCATGAAGAGAAATATGAAGAAATTTAAATCAGGATTAATGC 1465
Db 1559 TGTTAGGAGAGAGGCGCATGAAGAGAAATACGAGGAATTCAGGTAAGGATCAATGC 1618
QY 1466 CTTAGTTTCTAAGGCATTGAAAGTCCCGAGGAGGATGGATCATGCAAGTGGCACACC 1525
Db 1619 CTTGGTTGCAAAAGCCCAAGAGTTCTCTGAGGAAGATGGACAATCAAGATGGAAGCCC 1678
QY 1526 ATGCCAGGAAACAATACAGGAGTATCTGGAATGATTCAGTTTTCCTTGGTCACAG 1585
Db 1679 CTGGCTTGGAAACAACAGTACGCCGATCATCTCTGGAATGATTCAGGTAATCTTGGCCAAAG 1738
QY 1586 TGTGGCCTTGTATCTACGAGGTAAATGAGCTCCCGCTTTAGTTTATGTCTCGTGAAAA 1645
Db 1739 TGGCGGTGCTGATGTGGAAGGAATGAGTTGCCCTCGCTTGTATGCTCGAGAGAAA 1798

QY 1646 GCGTCTCGGTTCCAGCACCAAGAGGCTGGTCCCATGAATGCCCTTGTTCGTCTCTC 1705
Db 1799 GAGGCCAGGTTATTAACCATCAAGAGGCTGGTCCCATGAATGCATGGTCCGTGTCTC 1858
QY 1706 AGCTGTCTTACTAATATGGACAATACATGTTGAATCTTTGATTTGTGATCACTACATCAACAA 1765
Db 1859 TGCTGTCTTATCAAAATGCTGCATACCTATTGAACCTGGACTGTGATCACTACATCAACAA 1918
QY 1766 CAGCAAGGCTGTCCGAGAAAGCTATGTCTTCCCTAATGGATCCAAACCTAGTCTCCGACGT 1825
Db 1919 TAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCCTTTGGTGGGAAAGAAAGT 1978
QY 1826 CTGTTATGTGAGTTTCCCAAAAGGTTTGATGGGATTCATAGGAATGATCATATGCAAA 1885
Db 1979 GTGCTATGTACAGTTCCTCAGAGGTTGATGTTGATGACAAAATGATCATAGCTAA 2038
QY 1886 CAGGAACACTGTCTTTTTTGTATTTAACTTGAGGGGCTTCAGCGCATTCGAAGACCACT 1945
Db 2039 CAGGAACCTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGCTATTCAAGGCCAT 2098
QY 1946 TTATGTGGAACTGGTGTGTTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAAT 2005
Db 2099 TTATGTGGTACTGGATGTGTTTTCAGACGGCAGGCACTGTATGTTATGATGCTCTTAA 2158
QY 2006 TAAAGCGAAGAACCA-----GGTTTCTTTGGCATCACT 2038
Db 2159 AAGCAAGAACGCCACCTCAAGAACTTGCAACTGTGGCCCAAGTGGTGCTCTCTTGTGTG 2218
QY 2039 ATGTGGGGCAGAAAGAGGCAAGCAAGTCAAAGAAAGAGAGCTCAGATAAGAAAAA--- 2095
Db 2219 CTCGACGAGGAACAAGATAAAAGAACACTACAAAACAAAGAGGAGAGAAAG 2278
QY 2096 ---GTCGAACAAGCATGTGGACAGTTCGTTCCAGTATTTCAATCTCGAAGACATAGAGGA 2152
Db 2279 ATTATTTTCAAGAAAGCAGAAACCCATCTCTGCATATCTTTGGGTGAATTTGATGA 2338
QY 2153 GGGTCTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGCTCAAAATGAGCTT 2212
Db 2339 AGGTCTCCAGGTGC-----TGATATCGAAGAGCCGGAATCGTAATCAACAGAACT 2392
QY 2213 AGAAGAGATTTGGCCAGTCAGCAGCATTTTGGCTTCCACTCTGTATGGAAATATGGTG 2272
Db 2393 AGAAGAAATTTGGGCAGTCTTCTGTTTGTGTCATCAACACTTCTTGAGAACGGAGG 2452
QY 2273 TGTTCCTCAGTCTCCATCCAGAAATCTCTTTTGAAGAAAGCTATCCATGTCATAAGTTG 2332
Db 2453 GACCCCTGAAGAGCGCAAGTCCAGCTTCTTCTTGAAGGAAGCTATACATGTTATCAGCTG 2512
QY 2333 TGGCTATGAGGACAAAGTCTGAATGGGAACTCAGATTGGTTGGATCTATGATCTGTAC 2392
Db 2513 CGGCTACGAAGACAAAGACCGACTGGGGAAGAGAGATTGGGTGGATTTACGGATCGATC 2572
QY 2393 AGAAGATATTTCTACTGGATTCAAGATGCAGCAAGAGGCTGGGCTTCAGTCTATTGCA 2452
Db 2573 AGAGATATCTGACTGGATTTAAGATGCATGCCATGGCTGGCGGTCTATTTACTGCAT 2632
QY 2453 GCCCAAGGCCAGCTTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATCTCTGGAACA 2512
Db 2633 CCCGAAGCGGCTGCTATTCAAAGGTTCTGCGCTCTGAAACCTTTCGACCCGCTCTCACCA 2692
QY 2513 AGTGTCTGGTGGGCTCTCGGTTCTGTTGAAATCT--TTTCAGCCGGCATTTGCCCTTAT 2571
Db 2693 GGTCTCTCGTGGGCTTGGGCTCGGTGGAATTTTCTTCAGNANCAATGGNCACTTT 2752
QY 2572 GGTATGGCTACGAGGCGGCTCAAGTTTCTTGGAGAGATTCGCTTACATCAACACCA 2631
Db 2753 GGTACGGATACGCGCGGCTAAATTCCTGGAAAGGTTTCTTATATCAACTCCATCG 2812
QY 2632 TTTTACCCACTAACCTCTCTCCCGCTTCTAGTCTATTTGATATATGCTGCTATCTGTCTGC 2691
Db 2813 TTTATCCCTGGAGCTCCATCTCTCTCTGCTTACTGTACCTTGCTGCTGCCATCTGCCCTGC 2872

Qy	2692	TCACTGGAAAGCTTATCATCGCCAGAGATTACAACTGGCCAGTATCTGGTTTCATGGCG	2751
Db	2873	TCACGGGGAAGTTTATACACACGAGCTTACCAATGTCGCCAGTATCTGGTTCATGGCAC	2932
Qy	2752	TCCTTCTTTCAATTTTCGCCACTGGTATCCTTGAGATGAGGTGGAGTGGTGTGGCATTTG	2811
Db	2933	TTTTCATCTGCATCTCCGTGACCGGCATCCTGGAAATGAGGTGGAGTGGCGTGCCATCG	2992
Qy	2812	ACGAGTGGTGAGGAATAACAGTTCCTGGGTTCATTTGGAGTATCTCTCCACATCTGTTTG	2871
Db	2993	ACGACTGGTGAGGAACAGCAGTCTTGGGTTCATCGGGAGGCGTTTCGGCGCATCTGTTTCG	3052
Qy	2872	CCGTCTTTCAGGGTCTTCTGTAAGTGTCTGCCGTTATCGACACCAACTTCACTGTCACT	2931
Db	3053	CGGTGTTCCAGGGCGCTGCTGAAGTGTTCGCCGGCATCGACACGAGCTTCACCGTGACGT	3112
Qy	2932	CAAAAGGCTAATGACGAAGAAGCGACTTGTCTGAGCTCTACATGTTCAAGTGGAGGACGC	2991
Db	3113	CGAAGGCCGGGACGACGA---GGAGTCTCGGAGCTCTACAGTTCNAAGTGGACACCC	3169
Qy	2992	TTCTCATCCCTCCGACGACCAATTTTGATCATTAACATGGTTGGTGTCTGCTGGCACCT	3051
Db	3170	TGCTGATACCCCCACACACGCTCCTCTGCTGTAACCTTCAATCGGGTGGTGCCGGGATCT	3229
Qy	3052	CCTACGGCATCAACAGTGGTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTTCTTTG	3111
Db	3230	CGAAGCGCATCAACACGGGTACGAGTCTGGGGCCCCCTGTTCCGGNAAGCTCTTCTTCG	3289
Qy	3112	CCTTCTGGGTGATTGTTCACTTATACCCATTCCTCAAGGGTCTTATGGGCAAGGCAAAACC	3171
Db	3290	CGTTCTGGGTGATCGTCCACCTGTACCCGTTCTTCAAGGGTCTGGTGGGAGGAGACA	3349
Qy	3172	GCACACCGACGATTGTTCATCGTCTGGGTGCTCCTCCCTGCTTCTATCTTCTCCTTGGTGT	3231
Db	3350	GGACGCCGACGATCGTCATCGTCTGGTCCATCCTCTGGCTGGCTCGATCTTCTCGCTCTGT	3409
Qy	3232	GGTTCGTGTGTATCCATTCACTACCCGTCCTCGCTGGCCAAATATCAAAACCTGTGGCA	3291
Db	3410	GGTTCGGGTGCGACCCGTTCTTCGCCAAGACGACGGCCCGCTCCTGGAGGAGTGTGCC	3469
Qy	3292	TCAACTGCTAGGAAGTGGGAG	3313
Db	3470	TGGACTGCACTGAAGTGGGG	3491

RESULT 7
 AY104730
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 3783)
 Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 Zea Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 2 (bases 1 to 3783)
 Unpublished (2002)
 Coe,E.C.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 Location/Qualifiers
 1..3783
 /organism="Zea mays"
 /db_xref="MaizeDB:635765"
 /db_xref="taxon:4577"

Qy	1247	CGAGCTTCAATGCTGACHTTTTGACGCAATTGGCTGAGACTTCAGAGTTTGTCTAGGAATG	1306
Db	1334	TGTTGTCGCAATGCTAACGTTTTGAAGCAATTAICTGAAACATCTGAAATTTGCAAGAATG	1393
Qy	1307	GGTACCAATTGTGAAGAAGTATGACATTGAACCCAGAGCTCCCGAGTTTACTTTTGC	1366
Db	1394	GGTTCCTTCTGCAACGCTACAATATTTGAACCTCGCCTCCAGATGGTACTTCCAAACA	1453
Qy	1367	GAAAAATTGATTTACCTGAAAGACAAAGTCCAGCCTTCAPTTGTTTAAAGACCGCGGGCCAT	1426
Db	1454	GAAGATAGACTACTTTGAAAGACAAGGTCGACAGCAAACTTTGTTAGGAGAGGAGAGCAAT	1513
Qy	1427	GAAGACAGAAATATGAAGAAATTTAAATCAGAGATAATGCCCTAGCTTCTTAAGCATTTGA	1486
Db	1514	GAAGAGAGAGTATGAGGAATTCAGGTGAGAACTCAATGCCCTTAGTGTCCAAAGCCCGAA	1573
Qy	1487	AGTCCCGAGGAGGATGGATCATCAAGATGGCACACCATGGCCAGGAACAACATACAG	1546
Db	1574	AGTTCCTGAAGAAGGATGACAAATCAAGATGGAACCCCTGGCTGGAAACAATGTTG	1633
Qy	1547	GGATCATCTCGAATGATTCAGGTTTTCCTTGGTCCACAGTGGTGGCCTTGATCTAGGG	1606
Db	1634	TGATCATCTCGGAATGATTCAGGTTCTCCTTGGCCAAAGCGGAGGCCCTTGACTGTGAGG	1693
Qy	1607	TAATGAGCTCCCGCTTTAGTTTATGTGTCGTGAAAGCGTCTGGTTCAGACACCA	1666
Db	1694	AAATGAATGCCAGGATGGTTTATGTTTCTAGAGAGAAACGACCGCTATTAACATCA	1753
Qy	1667	CAAGAAGCGTGGTCCCATGAATGCCCTTGTGTGTCAGTCTGCTTACTTAATGGACA	1726
Db	1754	TAAGAAGCTGGTGTCTATGAATGATGTCGCGAGTCTGCTGTACTACAACATGCTCC	1813
Qy	1727	ATACATGTTGAATCTTGATTGTGATCACTACATCAACAACAGCAAGCTGTCCGAGAAGC	1786
Db	1814	ATATTTGTTAACTTGGATTGTGATCACTACATCAACAACAGCAAGGCTATTAAGGAGC	1873
Qy	1787	TATGTGCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTTATGTGCAGTTTCCACA	1846
Db	1874	AATGTGTTTTATGATGGACCCCTTACTAGAAAGAAAGGTTTGGCTATGTACAGTTCCCTCA	1933
Qy	1847	AAGTTTGTATGGGATGATGAGAAATGATCGATATGCAACAACAGGAACACTGCTCTTTTGA	1906
Db	1934	AGAATTTGATGGGATGATCGCCATGACCGATATGCTTAACCGGAATGTTGTCTTTTGA	1993
Qy	1907	TATTAACCTGAGGGCCTTGACGGCATTCAGGACCACTTATTTGTGGAACTGGTGTGT	1966
Db	1994	TATCAACATGAAGGTTTGGATGGTATTCAAGGTCCTCAATTTATGTTGGTACTGGATGTT	2053
Qy	1967	TTTCAACAGACAGCTATCTATGTTTATGACCCGCCCAATTAAGCGGAGAGCCA	2021
Db	2054	ATTTAGAGCGAGCATTAATGGTTATGATGCCCCCAACAAAGAACGCCACCATCAAG	2113
Qy	2022	-----GGTTTCTTGGCATCACTATGTGGGGCGCAAGAGAGGC	2059
Db	2114	GACTTGCAACTGCTGGCCCAAGTGTGCTTTTGTGTTGCTGCTTGGCAATAGGAACA	2173
Qy	2060	AACCAAGTCAAGAAAGAGCTCAGATAAGAAAAG-----TCGAACAGACGATGTGGA	2113
Db	2174	AAAGAAGCTACCAAAACCCAAACAGAGAAGAAAAGTTATTTATTTCAAGAAAGAAGA	2233
Qy	2114	CAGTTCTGTTCAGTATTCAATCTCGAAGACATAGAGAGGGTGTGAAGGTGCTGGGTT	2173
Db	2234	GAACCAATCCCCTGCATATGCTCTTGGTGAATTTGACGAAGTCTGCTCCAGGAGC-----	2287
Qy	2174	TGATGATGAGAAATCAGTTTCTCATGTCTCAATGAGCTTGTAGAGAGAGATTTGGCCAGTC	2233
Db	2288	TGAGAATGAAAGGCCGATTTGTAATCAACAAAAATTAGAAAAGAAATTTGGCCCAATC	2347
Qy	2234	ACAGCATTTGTTGCTCCACTCTGATGGAATATGGTGGTCTTCTCAGTCTCCACTCC	2293
Db	2348	TTCTGTTTTTTTACATCCACACTTCTCGAAGTGGTGGAACTTGAAGAGTGTCAAGTCC	2407
Qy	2294	AGAATCTCTTTTGAAGAAGCTATCCATGTCATAGTTTGTGGCTATGAGGACAGTCTGA	2353

D	b	2408	TGCTTCTCTTTTGAAGAAAGCACTATACATGTCTAATTAGTTGGTTATGAAGAACAAGACAGA	2467
Q	y	2354	ATGGGAAACTGCAGATTGGATTGGATCTCATGGATCTCTCACAGAAGATAATTCCTTACTTGCAAT	2413
D	b	2468	CTGGGNAAGAGATTGGCTGGATCTATGGATCAGTTACAGANAATATCTTAACCTGGTTTT	2527
Q	y	2414	CAGATGACGCGCAAGAGGCTGGGTTTCAGTCTATTGTCATGCCAACGGCCACGCTTTCAA	2473
D	b	2528	CAAGATGCATGTGCATGGTTGGCGGTCAATTTACTTGCATACCTTAACGGGTTGCATTTCAA	2587
Q	y	2474	GGATCTGCCCCCACATCAATCTTTCAGATCGCTGCTGAACCAAGTCTGGGGTGGGCTCTCGG	2533
D	b	2588	AGGTTCTGCACCTCTGAATCTTTTCAGATCGCTTCACCAGGTGCTTCGGTGGGCTCTTGG	2647
Q	y	2534	TTCTGTTGAAATTTCTTTCAGCGGCAATGGCCCCCTTATGGTATGGCTACGAGAGGGCGCCT	2593
D	b	2648	GTCATTTAGATCTCTTTCAGCAATCATTTGCCCTCTTTTGGTATGGGTATGGTGGCGGTCT	2707
Q	y	2594	CAAGTTCCTCGAGAGATTGCGTTACATCAACACACCAATTTACCACCTAACCTCTCTCCC	2653
D	b	2708	GAAATTTTGGAAAGATTTTCCCTACATCAACTCCATCGTGTATCCTTGGAGATCTATTCC	2767
Q	y	2654	GCCTTAGTCTATTGTATATTCGCTGCTATCTGTCTGCACCTGGAAGTTCAATCATGCC	2713
D	b	2768	CTCTTGGCTTACTGTACATTGCCCTGCCATCTGTTATTGACAGGAAATTTATCACTCC	2827
Q	y	2714	AGAGATTAGCAACTTGGCCAGTATCTGGTTCAATCGCGCTCTTCTTCCAATTTTCGGCCAC	2773
D	b	2828	AGAGCTGAATTAATTTTGGCAGCGCTGGTPTCATGTCACTTTTTTAFTGCAATTTTGGTAC	2887
Q	y	2774	TGGTATCTCTTGACATGAGGTGGAGTGGTCTTGGCAATTCACGAGTGGTGGAGGAATGAACA	2833
D	b	2888	GAGCATCTAGAATCAGATGGAGTGGTCTTGGAAATTCATGACTGGTGGAGAAATGAGCA	2947
Q	y	2834	GTTCCTGGGTCATTTGGAGGTATCTCTGCACATCTGTTTGGCGTCTTTTCAGGGTCTTCTGAA	2893
D	b	2948	GTTCTGGGTCATTTGGAGGTGTGCTCTCACACCTCTTGTCTGTGTTCAGGGACATCTCAA	3007
Q	y	2894	GGTCTTGGCGGTATCGACACCAACTTCACTGTCACTCAAGGCTAATGACGAAGAAGG	2953
D	b	3008	GGTCATAGCTGGTGATGATACAAGCTTCACCTGGACATCAAGG---GTGGAGATGATGA	3064
Q	y	2954	CGACTTTGTAGCTCTACATGTTCAAGTGGACGAGGTTCTCAATCCCTCCGAGACCAT	3013
D	b	3065	GGAGTTCTCAGAGCTATATACATTCAATTFSGACTACCTTATTGATACCTCTCTACACCTT	3124
Q	y	3014	TTTGATCATTAACATGTTGGTGTGCTTCTGCTGSCACCTTCTACGCCATCAACAGTGTGA	3073
D	b	3125	GCITCTATTGAATCTCATTTGGTGTGGTGTGCTGCGCTTCAAATGGGATCAATAACGATGA	3184
Q	y	3074	CCAAATCATGGGGCGGCTCTTTGGGAAGCTCTTCTTGGCTCTCTGGGTGATTGTTCACCT	3133
D	b	3185	TGAGTCAATGGGGCCCCCTCTTTGGGAAGCTATTCTTTGCAATTTTGGGTGATTGTCCATCT	3244
Q	y	3134	ATACCAATTCCTCAAGGTTCTTATGGCAGGCAAAACCGCACCCGACGAGTATGTCATCGT	3193
D	b	3245	TTATPCCCTTTCTCAAAAGGTTTGGTTGGAAAGGCAAAACAGACACCAAGATTTGTCATCGT	3304
Q	y	3194	CTGGGCTGTCTCTCTCGCTCTATCTTCTTCCCTTGTCTGTGGGTTGGTGTGTATCCATTCAC	3253
D	b	3305	CTGGTCCATTTCTGCTGCTTCAATCTCTCTGCTCTCTTTGGGTCGGATTGATCTCTTCTCT	3364
Q	y	3254	TACCGTCTCGCTGGGCCCAAAATATCAAAACCTGTGGCATCAACTGCTAGGAAAATGGGAG	3313
D	b	3365	TGCGAAAGGATGATGGTCCGCTTCTTGAGGAGTGTGGTTTGGATTGGCACTAGGATGTCAG	3424

RESULT 9

BM816138	BM816138	880 bp	mRNA	linear	EST 05-MAR-2002
LOCUS	HC109B12_T3.ab1	HC Hordeum vulgare	cdna clone	HC109B12_T3.ab1	
DEFINITION	similar to (AF200528) cellulose synthase-4 [Zea mays], (AF200529) cellulose synthase-5 [Zea mays], (AF200533) cellulose synthase-9 [Zea mays], cellulose synthase catalytic subunit [Arabidopsis thaliana], unnamed protein product [Arabidopsis thaliana], mRNA				
ACCESSION	BM816138	BM816138	1	GI:19152152	
VERSION	BM816138	BM816138	1	GI:19152152	
KEYWORDS	EST.				
SOURCE	Hordeum vulgare.				
ORGANISM	Hordeum vulgare.				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.				
AUTHORS	1 (bases 1 to 880)				
TITLE	Ozturk,N.Z., Michalowski,C.B., Brazille,S., Borchert,C., Palacio,C., Normand,C., Murphy,C., Kelley,R., Sant,S.A., McLaughlin,H., Fredricksen,M.A. and Bohnert,H.J.				
JOURNAL	Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley				
COMMENT	Unpublished (2002)				
FEATURES	Contact: Mark A. Fredricksen Plant Biology University of Illinois 1201 W Gregory Dr, Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu. Location/Qualifiers 1..880				
source	/organism="Hordeum vulgare" /strain="cv tokak" /db_xref="taxon:4513" /clone="HC109B12_T3.ab1" /clone_lib="HC" /tissue_type="Root" /dev_stage="3 week old" /note="6 and 10 hour drought stress by placing plants on moist paper (75% rel. humidity) in light"				
BASE COUNT	166 a	240 c	209 g	263 t	2 others
ORIGIN	Query Match Best Local Similarity 96.2%; Pred. No. 1.1e-201; Matches 801; Conservative 0; Mismatches 30; Indels 2; Gaps 2;				
QY	2469	TTCAAGGATCTGCCCAATCTTTCAGATCGTGAACCAAGTGTGCGTGGGCT	2528		
Db	48	TTTCAAGGATCTGCCCAATCTTTCAGATCGTGAACCAAGTGTGCGTGGGCT	106		
QY	2529	CTCGGTCTGTGAAATCTTTTCAGCGGCATTTGCCCTTATGGTATGGCTACGGAGGG	2588		
Db	107	CTTGGTTCGCTGAATCTTTTCAGCGGCATTTGCCCTTATGGTATGGCTACGGAGGG	166		
QY	2589	CGCCTCAAGTCTCGGAGATTCGGTTACATCAACACCACTTACCCACTAACCTCT	2648		
Db	167	CGCCTCAAGTCTCGGAGATTCGGTTACATCAACACCACTTACCCACTAACCTCT	226		
QY	2649	CTCCCGCTCTAGCTATTGATATTCCTGCTATCTGCTGCTCACTGGAAGTTTCATC	2708		
Db	227	CTCCCGCTCTAGCTATTGATATTCCTGCTATCTGCTGCTCACTGGAAGTTTCATC	286		
QY	2709	ATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTCGCTCTCTTCAATTTTC	2768		
Db	287	ATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTCGCTCTCTTCAATTTTC	346		
QY	2769	GCACATGGTATCTCTGAGATGGAGTGGTGGTGGCATTCACAGTGGTGGAGAAAT	2828		
Db	347	GCACCGGTATCTCTGAGATGGAGTGGTGGTGGCATTCACAGTGGTGGAGAAAT	406		
QY	2829	GAACAGTCTCGGTTCATTTGAGGTATCTCTGCACATCTGTTCCGCTTTTCAGGCTTT	2888		
base count	212 a	209 c	212 a	209 c	212 a 209 c

Db	407	GAACAGTCTCGGTTCATTTGGGCTCTTTTCAGGGTCTT	466	CTCTGCCCATCTGTTTGGCGTCTTTTCAGGGTCTT
QY	2889	CTGAAGTGTCTTGGCGGTATCT	2948	CAACTTCACTGTCACTCAAAAGGCTTAATGACGAA
Db	467	CTGAAGTGTCTTGGCGGTATCT	526	CAACTTCACTGTCACTCAAAAGGCTTAATGACGAA
QY	2949	GAAGGCGACTTTGTGTAGCTCT	3008	GTTCAGTGGAGGAGGCTTCTCATCCCTCCGACG
Db	527	GAAGGCGACTTTGTGTAGCTCT	586	GTTCAGTGGAGGAGGCTTCTCATCCCTCCGACG
QY	3009	ACCATTTTGTATTAATATG	3068	TGTGCTGCTGGCACCTCTTACGCCATCAACAGT
Db	587	ACCATTTTGTATTAATATG	646	TGTGCTGCTGGCACCTCTTACGCCATCAACAGT
QY	3069	GGTTACCAATCATGGGGCGG	3128	TGGGAAGCTCTTCTTGGCTCTTCTGGGTGATTGTT
Db	647	GGTTACCAATCATGGGGCGG	706	TGGGAAGCTCTTCTTGGCTCTTCTGGGTGATTGTT
QY	3129	CACATATACCATTCCTCAAG	3188	TATGGGAGGCAAAACCGCACACCGACGATTGTC
Db	707	CACATATACCATTCCTCAAG	766	TATGGGAGGCAAAACCGCACACCGACGATTGTC
QY	3189	ATCGTCTGGGCTGTCTCTCT	3247	TATCTTCTCTGCTGCTGGGTCTGCTGT-TGATCC
Db	767	ATCGTCTGGGCTGTCTCTCT	826	TATCTTCTCTGCTGCTGGGTCTGCTGTGATCC
QY	3248	ATTCACTACCGCTGTCTCTCT	3300	TATCCAAACCTGTGGCATCAACTGCT
Db	827	ATTCACTACCGCTGTCTCTCT	879	TATATNCAACCTGTGGCATCAACTGCT
RESULT 10	BM816139 linear EST 05-MAR-2002			
LOCUS	BM816139			
DEFINITION	um vulgare cDNA clone HC105G09_T3.ab1 cellulose synthase-4 [Zea mays], (AF200529) cellulose synthase-5 [Zea mays], cellulose synthase-9 [Zea mays], cellulose synthase catalytic subunit [Arabidopsis thaliana], unnamed protein product [Arabidopsis thaliana], mRNA sequence.			
ACCESSION	BM816139			
VERSION	BM816139.1 GI:19152152			
KEYWORDS	EST.			
SOURCE	Hordeum vulgare.			
ORGANISM	Hordeum vulgare.			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.			
AUTHORS	1 (bases 1 to 916) Ozturk,N.Z., Michalowski,C.B., Brazille,S., Borchert,C., Palacio,C., Kelley,R., Sant,S.A., McLaughlin,H., Fredricksen,M.A. and Bohnert,H.J.			
TITLE	Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley			
JOURNAL	Unpublished (2002)			
COMMENT	Contact: Mark A. Fredricksen Plant Biology University of Illinois 1201 W Gregory Dr, Urbana, IL 61801, USA Tel: 2172655473 Email: bohnertlab@life.uiuc.edu. Location/Qualifiers 1..916			
FEATURES	/organism="Hordeum vulgare" /strain="cv tokak" /db_xref="taxon:4513" /clone_lib="HC" /tissue_type="Root" /dev_stage="3 week old" /note="6 and 10 hour drought stress by placing plants on moist paper (75% rel. humidity) in light"			
BASE COUNT	212 a	209 c	212 a	209 c

crystallization dishes. Five-day old seedlings were incubated at 50c for 2 days. Shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, and 60000 pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 212 a 148 c 195 g 195 t

ORIGIN

Query Match 19.0%; Score 687.8; DB 12; Length 750;
Best Local Similarity 95.6%; Pred. No. 7.2e-181;
Matches 718; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 1479 GCATTCAAGTCCCGAGGAAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAC 1538
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 GCATTGAAAGTCCCGAGGAAGGATGATCATGCAAGATGGCACACCATGGCCAGGAAC 60

QY 1539 AATACCAGGATCATCCTGGAATGATTCAGGTTTTCCTTGTCACAGTGGTGGCCTTGAT 1598
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 AATACCAGGATCATCCTGGAATGATTCAGGTTTTCCTTGTCACAGTGGTGGCCTTGAT 120

QY 1599 ACTGAGGGTAATGAGTCCCGTTAGTTTATGTGTCGTGAAAGGGTCTGGGTC 1658
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 ACCGAGGGTAATGAGTCCCGTTAGTTTATGTGTCGTGAAAGGGTCTGGGTC 180

QY 1659 CAGCACCAAGAAGCTGGTGCATGAATGCCCTTGTTCGTCTCAGCTGTCTTACT 1718
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 CAGCACCAAGAAGCTGGTGCATGAATGCCCTTGTTCGTCTCAGCTGTCTTACT 240

QY 1719 AATGGACAATACATGTTGAATCTTGATGTGATCACTACATCAACAACAGCAAGCTGTC 1778
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AATGGACAATACATGTTGAATCTTGATGTGATCACTACATCAACAACAGCAAGCTGTC 300

QY 1779 CGAGAAGCTATGTGCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAG 1838
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 CGGAAGCTATGTGCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAG 360

QY 1839 TTCCCAACAAGGTTTATGGGATTTAGGAATGATCGATATGCAACAGGAACACTGTC 1898
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TTCCCAACAAGGTTTATGGGATTTAGGAATGATCGATATGCAACAGGAACACTGTC 420

QY 1899 TTTTGTGATTAATTAACCTTGAAGGCTTGAAGGATTCAGGACCCAGTTTATGTGGAACT 1958
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 TTTTGTGATTAATTAACCTTGAAGGCTTGAAGGATTCAGGACCCAGTTTATGTGGAACT 480

QY 1959 GGTGTGTTTTCACAGACAGCTATCTATGTTATGAGCCCCCAATTAAAGCGAAGAAG 2018
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 GGTGTGTTTTCACAGACAGCTATCTATGTTATGAGCCCCCAATTAAAGCGAAGAAG 540

QY 2019 CCAGGTTTCTTGGCATCTATCTGGGGCAAGAAGCAAGCAAGTCAAGAAAGG 2078
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 CCAAGTTTCTTGGCATCTATCTGGGGCAAGAAGCAAGCAAGTCAAGAAAGG 600

QY 2079 AGCTCAGATAAGAAAAGTCGAACAGCATGTGGACAGTTCGTTCCAGTATTCATCTC 2138
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 AGCTCAGATAAGAAAAGTCGAACAGCATGTGGACAGTTCGTTCCAGTATTCATCTC 660

QY 2139 GAAGACATAGAGGAGGTGTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATG 2198
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 GAAGACATAGAGGAGGCGGTTGAAGGTGCTGGGTTTGATGATGAGAAAT- AATCTCATG 719

QY 2199 TCTCAATGAGCTTAGAGAGAGAGATTGGCC 2229
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
720 TCTCAATGAACTTTAAAGAAAAAATTTCCGCC 750

Search completed: February 25, 2003, 19:51:29
Job time : 4769 secs



Query Match 1.6%; Score 59; DB 1; Length 3624;
Best Local Similarity 54.4%; Pred. No. 6.6e-06;
Matches 141; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

[illegible]

APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Suttie, Janet L.
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 607518artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,504B
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/459,595
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
OTHER INFORMATION: optimized cryiB"
OTHER INFORMATION: /note= "Disclosed in Figure 6."

US-08-459-504B-6
Query Match 1.6%; Score 59; DB 3; Length 3624;
Best Local Similarity 54.4%; Pred. No. 6.6e-06;
Matches 141; Conservative 0; Mismatches 115; Indels 3; Gaps 1;
Qy 53 CCGCGCCATGACGGCGACGGCCCTGAAGTCGGGAGGACGGGCGGGGACGT 112
Db 3213 CCACGAGATCGAAGAACACCCGAGCTGAAGTTCAAGACCGGAGGAGGAGGT 3272
Qy 113 GTGCCAGATCTGCGCCGACGGCCTGGGCACACGCTTGGACGGCGACGTCTTCACCGCGCT 172

Db 3273 GTACC---CCACCGACACCGGCACCTGCAACGACTACACCGCCACCGAGGCGCCGCGG 3329
Qy 173 CGAGCTGCGCGCTTCCGGTCTGCGGCCCTGTAGAGCACCAGGAGGAGGAC 232
Db 3330 CTGCGCCGACGCTGCAACAGCCGCGGCTACGAGGACGCTACGAGGTGGAC 3389
Qy 233 CCAGGCTGCTCGAGTGAAGACCAAGTACAAGCGGCACAGAGGGGACCGGATCCG 292
Db 3390 CACGCGCAGCTGAACCTAAGCCACCTACGAGGAGAGACCTACACCGACGTGCGCG 3449
Qy 293 CGGGGAGGAAGCGGACGAC 311
Db 3450 CGACAACCACTGCGAGTAC 3468
RESULT 6
US-08-459-444-6
; Sequence 6, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
NUCLEIC ACID CODING SEQUENCE
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,444A
FILING DATE: 02-Jun-1995
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS

"Full-length, maize
CRYIB"
enclosed in Figure 6."

QY	53	CCGCGCCATGGAGGGGACGCG	Ed: 0.000 000	Mismatches 115;	Indels 3;	Gaps
Db	3213	CCACGAGATCGAGAACAACACG	Best Local Similarity 34.4%			
		AGCTGAAGTTTCAAGAACCGCAGGAGGAGGAGGT	Matches 141; Conservative			
		CCCTGAAGTCCGGGAGGACAGGGCGGGGACGT				

Qy	113	GTGCCAGATCTGGCCGACCGG - - SCAACACGTTGGAGCGGCAGCTTTCACCGCCTG	172
Db	3273	GTACC---CCACGCACACCGG - - SCAAGACTATACACGCCCCCACCAAGGGACACCGCGG	3329

Qy	173	CGACGTCGCCGTTC	CCCTGCTACGAGCAGCGCAAGGAGGCAC	232
Db	3330	CTGCGCGCACGCTTCAACAGC	ACGGCGGTACGAGCAGCGCTACGAGTTGGACAC	3389

DQ
233 CCAGGCTTGCCTCCAGTGGAAG
| ||||| | | ||||
AGTACAAGGCCACAGAGGGAGCCCAGCGATCCG 292

Db
3390 CACCGCCAGGTGAACTAAC
| ||||| | | ||||
CCTACGAGGAGGAGACCTACACCGAGGTGCCCG 3449

Q1 293 CGGGGAGGAGGCGCCAGAC 31
|| - - - |||| - -
Db 3450 CGACACCCACTGCGAGTAC 34

RESULT 8

US-09-547-422-6
; Sequence 6, Application US/095
; Patent No. 6320100

APPLICANT: Koziel, Michael
Desai, Nalin
Lewis, Kelly

;
;
;
;
;

; Evola, Steph
; Crossland, R

Wright, Mart

;
;
;
;

;
TITLE OF INVENTION:
SYN
TNS

NUMBER OF SEQUENCES: 94

CORRESPONDENCE ADDRESS: ADDRESSEE: No. 6322

STREET: 3054 Cornw
CITY: Research Tr

STATE: NC

COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM
MEDIUM TYPE: ELECTRONIC

COMPUTER: IBM PC

```

; OPERATING SYSTEM:
; SOFTWARE: Patent

```

CURRENT APPLICATION DATA

APPLICATION NUMBER
FILING DATE: 11-APR

```

;
CLASSIFICATION: <U>
;
PRIOR APPLICATION DATA:

```

APPLICATION NUMBER

FILING DATE: 02-JUN-2009
APPLICATION NUMBER:

; FILING DATE: 25-SEP-80
; APPLICATION NUMBER:

FILING DATE: 04-00

IC DNA SEQUENCE HAVING ENHANCED
CIDAL ACTIVITY IN MAIZE

ARTIS AGRIBUSINESS BIOTECHNOLOGY RESEARCH, INC.

artis Agribusiness Biotechnology Research, Inc.


```

DOS
1.0, Version #1.25
B,261B
7,205
1506/CIP3
ore 53.4; DB 1; Length
ed. No. 0.00072;
Mismatches 106; Index
CTCGAAGTCCGGGAGGCACGGGG
CAAGACTTTCTCGGACACGGCGG
EACCAGCTTTGGACGGCGGAGGTCT
CTCGAGCGGATCGGCGACGCCCC
TCCTTGCTACGAGCACGAGCGCA
||| ||| ||| ||| |||
PACCGCATCAACACGAGGGCG
STACAAGCGCCACAGAG 276
CAGAAGGTCCTCCGG 24748

```

Best Local Similarity 13.8%; Pred.: No. 0.00012;
Matches 67: Conservative 200: Mismatches 220: Indels 0: Gaps 0:

```

1506/CIP3

pre 53.4; DB 1; Length 28958;
ed. No. 0.00072;
Mismatches 106; Indels 0; Gaps 0;

CTCTGAAGTCGGGAGGACACGGGGCCGGGACGTG 113
|||||
TAAGACCTTCTCGGACACGCCGACGGCTACGGA 24585
|||||
TACACAGTTTGGACGGGAGCTGTTTACCGCCTGC 173
|||||
TCTCGAGCGGATCGGGGAGCGCCCTCGCCCGGAGA 24645
|||||
TCCCTGCTACGAGCACGAGCGCAAGGAGGGCACC 233
|||||
TACCGCCATCAACACGAGGGCGCTCGAGCGGT 24705
|||||
TACAGGCGCCACAGAG 276
|||||
TACAGAGGTCTCCGGG 24748

```

```

RESULT 12
US-08-456-837-6
; Sequence 6, Application US/08456837
; Patent No. 5643774
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight St
; APPLICANT: Ryals, John And
; APPLICANT: Gaffney, Thomas
; APPLICANT: Lam, Stephen F
; APPLICANT: Hammer, Phillip
; APPLICANT: Uknes, Scott Jo
; TITLE OF INVENTION: Genes
; TITLE OF INVENTION: antipa
; NUMBER OF SEQUENCES: 22

```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10532
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/456,837
;; FILING DATE: 01-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/457,205
;; FILING DATE: 01-JUN-1995
;; APPLICATION NUMBER: 08/258,261
;; FILING DATE: 08-Jun-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Elmer, James Scott
;; REGISTRATION NUMBER: 36,129
;; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8614
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28958 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-456-837-6

Query Match 1.5%; Score 53.4; DB 1; Length 28958;
Best Local Similarity 52.5%; Pred. No. 0.00072;
Matches 117; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 54 CGCGCATGGACGGCGACGGCGCTGCTGAAGTCCGGGAGGACGCGGCGCGGACGTG 113
Db 24526 CGCGCTTGGCGCCGCGCGCTCCAGACCTTCTCGNACACGCGCGCTACGGA 24585
QY 114 TGCCAGATCTCGCGCGACGGCTGGGACACACAGTTGGACGCGACGTCTTTCACCGCCTGC 173
Db 24586 CGCGGAGAAGCGTCTGCTGCTTTCGCTCGAGCGGATCGGCGACGCTCGCCCGGAGA 24645
QY 174 GACGCTCCGCTTCCGCGTCTGCGCCCTGCTACGAGCAGGACGCGCAAGGAGGCACC 233
Db 24646 CACCGCTCTCTGCTCTGCTCGCGGACCGCCATCAACACGACGCGCTCGAGCGGT 24705
QY 234 CAGGCGCTCCCTCCAGTGCAGACCAAGTACAAAGCCACAGAG 276
Db 24706 ATCACCAGCGCCCAACGGCACCTCCACAGAGAAGTCTCTCCGGG 24748

RESULT 13
US-08-457-342-6
; Sequence 6, Application US/08457342
; Patent No. 5662898
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.

;; APPLICANT: Uknes, Scott Joseph
;; TITLE OF INVENTION: Genes for the synthesis of
;; TITLE OF INVENTION: antipathogenic substances
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10532
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/457,342
;; FILING DATE: 01-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/457,205
;; FILING DATE: 01-JUN-1995
;; APPLICATION NUMBER: 08/258,261
;; FILING DATE: 08-Jun-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Elmer, James Scott
;; REGISTRATION NUMBER: 36,129
;; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8614
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28958 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-457-342-6

Query Match 1.5%; Score 53.4; DB 1; Length 28958;
Best Local Similarity 52.5%; Pred. No. 0.00072;
Matches 117; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 54 CGCGCATGGACGGCGACGGCGCTGCTGAAGTCCGGGAGGACGCGGCGCGGACGTG 113
Db 24526 CGCGCTTGGCGCCGCGCGCTCCAGACCTTCTCGGACACGCGCGCTACGGA 24585
QY 114 TGCCAGATCTGCGCGACGGCTGGGACACACAGTTGGAGCGGCTTTCACCGCCTGC 173
Db 24586 CGCGGAGAAGCGTCTGCTGCTTTCGCTCGAGCGGATCGCGGACGCTCTCCCGGAGA 24645
QY 174 GACGCTCCGCTTCCGCGTCTGCGCCCTGCTACGAGCAGGACGCGCAAGGAGGCACC 233
Db 24646 CACCGGCTCTCTGCTCTGCTCGCGGACCGCCATCAACACGACGCGCTCGAGCGGT 24705
QY 234 CAGGCGCTCCCTCCAGTGCAGACCAAGTACAAAGCCACAGAG 276
Db 24706 ATCACCAGCGCCCAACGGCACCTCCACAGAGAAGTCTCTCCGGG 24748

RESULT 14
US-08-457-646A-6
; Sequence 6, Application US/08457646A
; Patent No. 5679560
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven

Job time : 589 secs





GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 18:32:29 ; Search time 181 Seconds
(without alignments)
11251.406 Million cell updates/sec

Title: US-09-900-237-29
Perfect score: 3626
Sequence: 1 gcacgaggaacccgctcca.....tcacatttggaggagtttt 3626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 442118 seqs, 280819700 residues 884236

Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3626	100.0	3626	10	US-09-900-237-29
2	1681.6	46.4	1733	10	US-09-900-237-19
3	1364.4	37.6	3563	10	US-09-900-237-25
4	1340.2	37.0	3786	10	US-09-900-237-7
5	1320.4	36.4	3776	10	US-09-900-237-3
6	1271	35.1	3936	10	US-09-900-237-9
7	1257	34.7	3517	10	US-09-900-237-13
8	1215.6	33.5	3328	9	US-09-838-539-1
9	1192.6	32.9	3255	9	US-09-938-842A-559
10	1191.4	32.9	3255	9	US-09-938-842A-2530
11	1183.4	32.6	2890	10	US-09-900-237-11
12	1069.6	29.5	2663	10	US-09-900-237-23
13	966.8	26.7	2125	10	US-09-900-237-15
14	821.4	22.7	1560	10	US-09-900-237-27
15	524.8	14.5	1221	10	US-09-900-237-5
16	518.8	14.3	1189	10	US-09-900-237-5
17	419.6	11.6	1029	10	US-09-900-237-21
18	324.2	8.9	606	10	US-09-770-149-916
19	281.6	7.8	384	10	US-09-878-574-797

20	240	6.6	377	10	US-09-878-574-3023	Sequence 3023, Ap
21	202	5.6	283	10	US-09-923-876-4542	Sequence 4542, Ap
22	193.4	5.3	272	10	US-09-523-876-4807	Sequence 4807, Ap
23	187	5.2	271	10	US-09-878-574-11933	Sequence 11933, A
24	177.4	4.9	284	10	US-09-294-093B-1388	Sequence 1388, Ap
25	171.8	4.7	480	10	US-09-734-569-39	Sequence 39, Appl
26	171.8	4.7	1009	10	US-09-734-569-153	Sequence 153, App
27	167.2	4.6	255	10	US-09-878-574-5625	Sequence 5625, App
28	164.2	4.5	261	10	US-09-878-574-10197	Sequence 10197, A
29	160	4.4	278	10	US-09-294-093B-5812	Sequence 5812, Ap
30	157.8	4.4	280	10	US-09-878-574-14007	Sequence 14007, A
31	151.8	4.2	268	10	US-09-878-574-6337	Sequence 6337, Ap
32	149	4.1	259	10	US-09-878-574-13937	Sequence 13937, A
33	142.2	3.9	260	10	US-09-878-574-11087	Sequence 11087, A
34	138.6	3.8	496	10	US-09-734-569-37	Sequence 37, Appl
35	138	3.8	285	10	US-09-294-093B-4005	Sequence 4005, Ap
36	136.6	3.8	2268	10	US-09-887-576-669	Sequence 669, App
37	136.4	3.8	254	10	US-09-923-876-5625	Sequence 5625, Ap
38	134.6	3.7	258	10	US-09-923-876-2317	Sequence 2317, Ap
39	133.4	3.7	280	10	US-09-294-093B-1384	Sequence 1384, Ap
40	130	3.6	265	10	US-09-878-574-8805	Sequence 8805, Ap
41	128.4	3.5	474	10	US-09-770-444-100	Sequence 100, App
42	128.4	3.5	474	10	US-09-924-035A-217	Sequence 217, App
43	128	3.5	258	10	US-09-923-876-1909	Sequence 1909, Ap
44	127.4	3.5	226	10	US-09-294-093B-4855	Sequence 4855, Ap
45	125.8	3.5	2139	9	US-09-938-842A-2652	Sequence 2652, Ap

ALIGNMENTS

RESULT 1						
US-09-900-237-29						
; Sequence 29, Application US/09900237						
; Patent No. US20020120124A1						
; GENERAL INFORMATION:						
; APPLICANT: Allen, Stephen						
; TITLE OF INVENTION: Plant Cellulose Synthases						
; FILE REFERENCE: B81170 US CIP						
; CURRENT APPLICATION NUMBER: US/09/900,237						
; PRIOR FILING DATE: 2001-07-06						
; PRIOR APPLICATION NUMBER: 60/092,844						
; PRIOR FILING DATE: 1998-07-14						
; PRIOR APPLICATION NUMBER: PCT/US99/15871						
; PRIOR FILING DATE: 1999-07-13						
; PRIOR APPLICATION NUMBER: 09/720383						
; PRIOR FILING DATE: 2000-12-21						
; NUMBER OF SEQ ID NOS: 33						
; SOFTWARE: Microsoft Office 97						
; SEQ ID NO 29						
; LENGTH: 3626						
; TYPE: DNA						
; ORGANISM: Triticum aestivum						
US-09-900-237-29						
Query Match						
Best Local Similarity 100.0%; Score 3626; DB 10; Length 3626;						
Matches 3626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	GCACGAGGAACCCGCTCCAGCTCTGTCGTCGGTGGGTTGGATFCGCTCTGCCGCGCCA	60			
Db	1	GCACGAGGAACCCGCTCCAGCTCTGTCGTCGGTGGGTTGGATFCGCTCTGCCGCGCCA	60			
QY	61	TGGAGCGGACCGCGACCCCTGAGTCCGGAGGACGCGGGCGGACCTGTGCCAGA	120			
Db	61	TGGAGCGGACCGCGACCCCTGAGTCCGGAGGACGCGGGCGGACCTGTGCCAGA	120			
QY	121	TCTGCGCGACGCGCTGGCGACCCAGCTTGGACGGGACGCTTACCCGCTGCCACGCT	180			
Db	121	TCTGCGCGACGCGCTGGCGACCCAGCTTGGACGGGACGCTTACCCGCTGCCACGCT	180			
QY	181	GCGGTTCCCGGTCGTGCGGCCCTCTACGACGACGCGCAAGGAGGCGCACCGGCT	240			

Db 181 GCGGCTTCCGGTGTGCGGCCCTGCTACGAGCAGCAGCGCAAGGAGGCGCACCGCCCT 240
QY 241 GCCTCCAGTCAAGCAAGTACAAGCGCCACAGAGGAGGCCAGCGCATCCGGGGAGG 300
Db 241 GCGTCCAGTCAAGCAAGTACAAGCGCCACAGAGGAGGCCAGCGCATCCGGGGAGG 300
QY 301 AAGCGACGACACTGATGCGCGATGATGGTAGTGACTTCAACTACCCCTGCACTCTGGCACTG 360
Db 301 AAGCGACGACACTGATGCGCGATGATGGTAGTGACTTCAACTACCCCTGCACTCTGGCACTG 360
QY 361 AGGACCAAGACAGAGAGTAGTCTGACAGATGCGCAGCTGGCGGATGAACACCGGGGGCA 420
Db 361 AGGACCAAGACAGAGAGTAGTCTGACAGATGCGCAGCTGGCGGATGAACACCGGGGGCA 420
QY 421 GTGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACA 480
Db 421 GTGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACA 480
QY 481 GTGGAGAGATCCCTAGGGGATACGTCCTCTTCACTCAGTCAACCAAGCCAGATGTCAAGGAAA 540
Db 481 GTGGAGAGATCCCTAGGGGATACGTCCTCTTCACTCAGTCAACCAAGCCAGATGTCAAGGAAA 540
QY 541 TCCTTGAGCTTCGCCCTGATCATCATCATGATGTCCTTACGGGAGCATCAGCAGAGCTG 600
Db 541 TCCTTGAGCTTCGCCCTGATCATCATCATGATGTCCTTACGGGAGCATCAGCAGAGCTG 600
QY 601 CTCGGTTCCCTATGTGAATCATTCACCAATCCGTCAAGGGAGTTCTCGGGCAGTATG 660
Db 601 CTCGGTTCCCTATGTGAATCATTCACCAATCCGTCAAGGGAGTTCTCGGGCAGTATG 660
QY 661 GGAATGTTGGCTGGAAGAGAGAGTTGATGGCTGGAAAATGAAGCAGGAGGATGCGGA 720
Db 661 GGAATGTTGGCTGGAAGAGAGAGTTGATGGCTGGAAAATGAAGCAGGAGGATGCGGA 720
QY 721 TTCCCATGACTAATGGGACAGCATTCCTCCCTCTGAAGTGGGGCAGCTACTGACATCG 780
Db 721 TTCCCATGACTAATGGGACAGCATTCCTCCCTCTGAAGTGGGGCAGCTACTGACATCG 780
QY 781 ATGCATCTACTGAATCAACATGGAAGACGCTTTACTGAAATGATGAACCTCGCCAGCCTC 840
Db 781 ATGCATCTACTGAATCAACATGGAAGACGCTTTACTGAAATGATGAACCTCGCCAGCCTC 840
QY 841 TATCTAGAAAAGTCCCATTTGCTTCTCCAAAATAAATCCCTACAGAATGGTCAATGTTTC 900
Db 841 TATCTAGAAAAGTCCCATTTGCTTCTCCAAAATAAATCCCTACAGAATGGTCAATGTTTC 900
QY 901 TCGGTTGGTGTCTAAGCATCTTCTGCACCTACCCTCTCAAAAATCCCTGCGTAATG 960
Db 901 TCGGTTGGTGTCTAAGCATCTTCTGCACCTACCCTCTCAAAAATCCCTGCGTAATG 960
QY 961 CATACCCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTTATCTCTGGATAC 1020
Db 961 CATACCCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTTATCTCTGGATAC 1020
QY 1021 TGGATCAGTTCCGGAAGTGTTCCTCAATCAACCGGGAGACCTACCTTTGATAGACTGGCTT 1080
Db 1021 TGGATCAGTTCCGGAAGTGTTCCTCAATCAACCGGGAGACCTACCTTTGATAGACTGGCTT 1080
QY 1081 TAAGGTATGACCGAGAGGTGAACCGTCTCAGTTGGCTGTGTGACATATTTGTCAGTA 1140
Db 1081 TAAGGTATGACCGAGAGGTGAACCGTCTCAGTTGGCTGTGTGACATATTTGTCAGTA 1140
QY 1141 CAGTCGACCCCTTGAAGGAGCCACTATCGTCACTGCAACACTGTCTATCCATTTCTTG 1200
Db 1141 CAGTCGACCCCTTGAAGGAGCCACTATCGTCACTGCAACACTGTCTATCCATTTCTTG 1200
QY 1201 CTGTTGATATCCCGTGGACAAGTCTCTGCTATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 CTGTTGATATCCCGTGGACAAGTCTCTGCTATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 TGACTTTTGACGATGGCTGAGACTTCAGAGTTTGTAGAGAAATGGGTACCATTTGTGA 1320
Db 1261 TGACTTTTGACGATGGCTGAGACTTCAGAGTTTGTAGAGAAATGGGTACCATTTGTGA 1320

QY 1321 AGAAGTATGACATGAACCCAG 1380
Db 1321 AGAAGTATGACATGAACCCAG 1380
QY 1381 TGAAGACCAAGTCCAGCCTTC 1440
Db 1381 TGAAGACCAAGTCCAGCCTTC 1440
QY 1441 AAGAAATTAATCAGGATAA 1500
Db 1441 AAGAAATTAATCAGGATAA 1500
QY 1501 GATGGATCATGCAAGATGGCA 1560
Db 1501 GATGGATCATGCAAGATGGCA 1560
QY 1561 TGATTCAGGTTTCCCTTGGTCA 1620
Db 1561 TGATTCAGGTTTCCCTTGGTCA 1620
QY 1621 GTTTAGTTTATGTCTCTGTC 1680
Db 1621 GTTTAGTTTATGTCTCTGTC 1680
QY 1681 CCATGAATGCCCTTGTTCGTG 1740
Db 1681 CCATGAATGCCCTTGTTCGTG 1740
QY 1741 TTGATTTGATCACTACATCA 1800
Db 1741 TTGATTTGATCACTACATCA 1800
QY 1801 TGGATCCAAACCTAGGTCCGCA 1860
Db 1801 TGGATCCAAACCTAGGTCCGCA 1860
QY 1861 TTGATAGGAATGATCGATATG 1920
Db 1861 TTGATAGGAATGATCGATATG 1920
QY 1921 GCCTTGAGCGGATTCAGGAGC 1980
Db 1921 GCCTTGAGCGGATTCAGGAGC 1980
QY 1981 CTATCTATGTTATGACCCCG 2040
Db 1981 CTATCTATGTTATGACCCCG 2040
QY 2041 GTGGGGCAAGAGAAGGCAAG 2100
Db 2041 GTGGGGCAAGAGAAGGCAAG 2100
QY 2101 ACAAGATGTGGACAGTTCTG 2160
Db 2101 ACAAGATGTGGACAGTTCTG 2160
QY 2161 AAGTGTCTGGTTTGTATGATG 2220
Db 2161 AAGTGTCTGGTTTGTATGATG 2220
QY 2221 GATTTGGCCAGTCAAGCAGAT 2280
Db 2221 GATTTGGCCAGTCAAGCAGAT 2280
QY 2281 AGTCCTCCACTCCAGATCTCT 2340
Db 2281 AGTCCTCCACTCCAGATCTCT 2340
QY 2341 AGGACAAGTCTGAATGGGAG 2400
Db 2341 AGGACAAGTCTGAATGGGAG 2400

Db 421 CAATGACCTTAGAAGAGATTTGGCCAGTCACGCCATTTGTTGCCTCCACTCTGATG 480
QY 2262 GAATPATGTTGGTCTCCTCAGTCTCCTCCACTCCAGAAATCTCTTTTGAAGAAGCTATCCAT 2321
Db 481 GAATATGTTGGTGTCTCCTCAGTCTCCTCAGTCCACATCCAGAAATCTCTTTTGAAGAAGCTATCCAT 540
QY 2322 GTCATAGTTTGGCTATGAGGACAAGTCTGAATGGGGAAGTCTGAGATTTGGTGGATCTAT 2381
Db 541 GTCATAAGTTTGGGCTATGAGGACAAGTCTGAATGGGGAAGTCTGAGATTTGGTGGATCTAT 600
QY 2382 GGATCTGTCCAGAAAGATATTTCTTACCTGGATTTCAAGATGCACGCAAGAGGCTGSCGTCA 2441
Db 601 GGATCTGTCCAGAAAGATATTTCTTACCTGGATTTCAAGATGCACGCAAGAGGCTGSCGTCA 660
QY 2442 GTCTATTGCATGCCAACGCCACAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGAT 2501
Db 661 ATCTATTGCATGCCAACGCCACAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGAT 720
QY 2502 CGTCTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTGAAATTTCTTTTCAGCGGCAT 2561
Db 721 CGTCTGAATCAAGTCTGCGGTGGGCTCTCGGTTCTGTTGAAATTTCTTTTCAGCGGCAT 780
QY 2562 TGCCCTCTATTGGTATGCTACGAGGGCGCTCAAGTTCTGGAGAGATTCGCTTACATC 2621
Db 781 TGCCCTCTATTGGTATGCTACGAGGGCGCTCAAGTTCTGGAGAGATTCGCTTACATC 840
QY 2622 AACACCACATTTACCACCAATCACTCTCTCCGCTTCTAGTCTATTTGATATTTGCCCTGCT 2681
Db 841 AACACCACATTTACCACCAATCACTCTCTCCGCTTCTAGTCTATTTGATATTTGCCCTGCT 900
QY 2682 ATCTGTCTGCTCACTGAAAGTTTCAATCATGCCAGAGATTTAGCAACTTTGGCCAGTATCTGG 2741
Db 901 ATCTGTCTGCTCACTGAAAGTTTCAATCATGCCAGAGATTTAGCAACTTTGGCCAGTATCTGG 960
QY 2742 TTCATTGCGCTCTCTCTTTTCAATTTTCCGCCATCGTATCTCTGAGATGAGTGGAGTGT 2801
Db 961 TTCATTGCGCTCTCTCTTTTCAATTTTCCGCCATCGTATCTCTGAGATGAGTGGAGTGT 1020
QY 2802 GTTGGCATTCACGAGTGGTGGAGGAATGAACAGTTCTTGGTCTATTTGGAGGTATCTCTGCA 2861
Db 1021 GTTGGCATTCACGAGTGGTGGAGGAATGAACAGTTCTTGGTCTATTTGGAGGTATCTCTGCC 1080
QY 2862 CATCTGTTTCCGCTCTTTCAGGCTCTTCTGAAGTGTCTTCCGGTATCGCGGTATCGACCAACTTC 2921
Db 1081 CATCTGTTTCCGCTCTTTCAGGCTCTTCTGAAGTGTCTTCCGGTATCGCGGTATCGACCAACTTC 1140
QY 2922 ACTGTCACTCAAGGCTAATGACGAAGAAGGCGACTTTGCTGAGCTTACATGTTTCAAG 2981
Db 1141 ACTGTCACTCAAGGCTAATGACGAAGAAGGCGACTTTGCTGAGCTTACATGTTTCAAG 1200
QY 2982 TGGACGACGCTTCTATCCCTCCGACGACCATTTTGTATCATTTAACATGTTGTTGTCGTT 3041
Db 1201 TGGACGACGCTTCTATCCCTCCGACGACCATTTTGTATCATTTAACATGTTGTTGTCGTT 1260
QY 3042 GCTGGACCTTCTACGCCATCAACAGTGTACCAATCATGCGGCGGCTCTTTTGGGAAG 3101
Db 1261 GCTGGTACCTTCTACGCCATCAACAGTGTGTACCAATCATGCGGCGGCTCTTTTGGGAAG 1320
QY 3102 CTCTTCTTTTGGCTTCTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGGTCTTATGGGC 3161
Db 1321 CTCTTCTTTTGGCTTCTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGGTCTTATGGGC 1380
QY 3162 AGGCAAAACCGCACACGACGATTTGTCATCGTCTGGGCTGTCTCTCGCTTCTATCTTC 3221
Db 1381 AGGCAAAACCGCACACGACGATTTGTCATCGTCTGGGCTGTCTCTCGCTTCTATCTTC 1440
QY 3222 TCCTTGTGTGGGTCTGCTGTTGATCCATTCACCTACCGCTCTCGTGGCCCAAAATATCCAA 3281
Db 1441 TCCTTGTGTGGGTCTGCTGTTGATCCATTCACCTACCGCTCTCGTGGCCCAAAATATCCAA 1500
QY 3282 ACCTGTGGCATCAACTGCTAGGAAAGTGGAGTTTGTAGAGACAGAAATATTAACAGTGA 3341
Db 1501 ACCTGTGGCATCAACTGCTAGGAAAGTGGAGTTTGTAGAGACAGAAATATTAACAGTGA 1560

QY 3342 TCGAGCAACACCCGCGGAGCG 3401
Db 1561 TCGAGCGCACCCCTGTGGAGCG 1620
QY 3402 GAGAAAGTTGTCAAAATTGAG 3461
Db 1621 GAGAAAGTTGTCAAAATTGAG 1680
QY 3462 TTTTCATGAGTTAAAGCTCTTC 1721
Db 1681 TTTTCATGAGTTAAAGCTCTTC 1780
RESULT 3
US-09-900-237-25
; Sequence 25, Application US/09-900-237-25
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cell Culture Media
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: 09/000,237
; PRIOR FILING DATE: 2001-07-14
; PRIOR APPLICATION NUMBER: 60/158,871
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US00/090,237
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/000,237
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 3563
; TYPE: DNA
; ORGANISM: Impatiens balsamifolia
US-09-900-237-25

Query Match 37.6%
Best Local Similarity 66.2%
Matches 2108; Conservative
Mismatches 1026; Indels 48; Gaps 8;

QY 81 CTGAAGTCCGGGAGGCGACGGG 140
Db 88 CCGAAGCCCTTAAAGAGGCTA 147
QY 141 ACCACGTTTGGACGCGGACGCT 200
Db 148 AATCAGCCACCGCGGACACTT 207
QY 201 CCCTGTACGAGCACGACGCGCA 260
Db 208 CCTTGTACGAGTACGAAAGGA 267
QY 261 TACAAGCGCCACAGAGGGAGCG 320
Db 268 TACAAAAGACAGAAAGGAGTCT 327
QY 321 GATGATGGTGTAGTCACTTCACT 380
Db 328 GATTTGAAACACGATTTCAATTC 384
QY 381 GCTGACAGGATGCGGACGCTGCG 440
Db 385 ACGCCAAGCGGCGCATGGCAGG 444
QY 441 AAGTATGACAGTGGCGAGATCG 500
Db 445 AGCATGATGAATCCCAACAC 504
QY 501 TACGTCCCTTTCAGTTCACCAACA 560
Db 505 GAAATTC--CACTCCCGATAA 562

AAATATTATGTTGGGTTGTGAATTACTACGTTT 3401
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3401
AAATATTATGTTGGGTTGTGAATTACTACGTTT 1620
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1620
ACATTGTAAATAGATGTAATAGACCACCTACCG 3461
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3461
ACATTGTAAATAGATGTAATAGACTATCTACCG 1680
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1680
TTTGGAAACAAAGGAA 3502
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3502
TTGGAAAAAAAAAAAA 1721
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1721

QY 561 CATCATGATGTCCTTACGGGGAACATCAGCAGACGTGCTCCGTTCCCTTATGTGAAT 620
Db 563 TGGAAAAATCCATTCCTCTATATCGATCCAGGACGACGATTCGTGAGAAATATTTGTG 622
QY 621 CATTCACCAAAATCCGTCACAGGAGTTCTCCGGCAGTATTGGAAATGTTCCCTGGAAAGAG 680
Db 623 ACCCTTCCAAAGGACTTGAA-----CTCGTACGGGCTTGGCANAATGGAAGGAA 675
QY 681 AGATTGTAGTGGGTGGAATAAGCAGGACAAAGGTTGCGATTCCCATGACTAATGGGACA 740
Db 676 AGGGTTGAAGGTTGGAAACTTAAGCAGGAGGAAAAATATGGTGCAAAATGAC----- 725
QY 741 AGCATTCCTCCCTCTCAAGGTCGGCAGCTACTGACATCGATGCACTCTACTGAATACAC 800
Db 726 --CAGTAGATATCTCTGAAGGGAAAGGAGATCTGAAGGTACGGGATCAAAATGGGAGGAA 783
QY 801 ATGGAAGAGCGTTTACTGAATGATGAACCTCGCCAGCCCTCTATCTAGAAAAAGTCCCCATT 860
Db 784 CTTCAAAATGGCTGCC-----GACGATATCCGACACACCATGAGCCGAATCGTGCCTCAT 837
QY 861 GCTTCTCCAAAATAAATCCCTACAGAAATGGTCAATGTTCTCGGGTGGTTGTTCTTAAGC 920
Db 838 TCGTCAGCCACCTCACTCCCTACAGAGTGGTTATCATACTTCGGTTAATATCTCTCGGT 897
QY 921 ATCTTCTGCTACCTACCGTCTCAAAATCCTGTGCGTAAATGACATACCACCTGTGGCTTTA 980
Db 898 TTCTTCTTGAATACCGTTGTACTCATCCAGTGAAGATGCTTTATCCATTATGGCTTACC 957
QY 981 TCTGTTATATGTGAGATTTGGTTGCTTTATCCTGGATACTGGATCAGTTCCCGAAGTGG 1040
Db 958 TCGGTTATTTGTGAAGTTTGGTTTGCATTGTCATGSCCTGCTTGATCAGTTTCCCTTAATGG 1017
QY 1041 TTTCCAAATCAACGGGAGACCTACCTTGATAGACGTGGCTTTAAGGTATGACCGAGAAGT 1100
Db 1018 TCGCCCTGTGAACCGCGAGACTTATCTCGACAGACTGTCCATGAGATTCGATAGGAAGGG 1077
QY 1101 GAACCCGTCTCAGTGTGCTGCTGTGACATATTTGTGACGTACAGTCGACCCCTTTGAAGAG 1160
Db 1078 GAGCCCTCGCAATTGGCGCAATTGACGTATTTGTTAGTACCGTGGATGCTTTGAAAGAG 1137
QY 1161 CCACCTATCGTCACTGCCAACACTGTGCTATCCATTCTTGTGTTGATTATCCCGTGGAC 1220
Db 1138 CCACCACTGTGACAGCCAAACACGGTTTGTCTATCTCGCCGTGGATTACCCCTGTTGAC 1197
QY 1221 AAGGTCCTTGCTATGTAATGATGACGGAGCTTCAATGCTGACITTTGACCGCATGGCT 1280
Db 1198 AAGTCTCTTGTGATGTTTCGGATGACGGTTGAGCAATGTTGACCTTCGAAGCTCTATCC 1257
QY 1281 GAGACTTCAGAGTTTCTAGGAAATGGTACCATTGTTGAAGAAAGTATGACATTGAACCC 1340
Db 1258 GAGACAGCCGAGTTTCTGAAGAAATGGGCAACCTTCTGTGAAGAAATAGTATTGAACT 1317
QY 1341 AGAGCTCCCGAGTTTACTTTTCCAGAAATTTGATTACCTGAAAGACAAAGTCCAGCCT 1400
Db 1318 CGGCGCCCGAAATTTTATTTCCGTCAAAAGATTGATTACTTGAAGGATAAGGTGACGCT 1377
QY 1401 TCATTGTTTAAAGACCGCCGGGCCATGAAGAGAGAAATATGAAGAAATTTAAATTCAGGATA 1460
Db 1378 TCTTTCTGTGAAGAGCGGAGGCGGATGAAGAGGGAATACGAAGAATTTAAGGTTAGGATT 1437
QY 1461 AATGCCCTAGTTTCTAAGGCATTGAAAGTCCCGAGGAGGATGGATCATGCAAGATGCC 1520
Db 1438 AATGCCCTGTTGCGGAAGCGGCAAAAGTGCAGAAAGAGGATGGACGATGCAAGATGGA 1497
QY 1521 ACACATGGCCAGGAAACAAATACAGGAGTCACTCTCGGAATGATTCAGGTTTTCTCTGGT 1580
Db 1498 ACTCCATGGCCGGAATAACTCGAGAGATCATCTCGGAATGATTCAGGTTTTTTTAGCC 1557
QY 1581 CACAGTGGTGGCCTTGATACTGAGGGTAATAGCTCCCGGTTTAGTTTATGTGCTCGT 1640
Db 1558 CATAGTGGGGGTTTCGATACGAGGGAATGAGTTACCTCGGCTGGTGTACGTTTCTCGT 1617

QY 1641 GAAAGCGCTCTCGGTTCCAGCACCAAGAGAGCTGGTGCCATGAATGCCCTTGTTCGT 1700
Db 1618 GAGAAACGCTCTGGATTTCAGCATCACAAAGAACCGGGGCAATGAACGCATTGATTCGA 1677
QY 1701 GTCTCAGCTGCTCTTACTTACTTGGCAATACATCTTGAATCTTGAATCTTGTGATCACTACATC 1760
Db 1678 GTATCGGAGTGTCTGACAAATGGGGCTTATCTGCTTAACGTGGATTGTGATCACTACTTTC 1737
QY 1761 AACACAGCAAGGCTGTCGGAAGAGCTATGTGCTTCCCTAATGATGATCAAAAGTCTAGGTCCG 1820
Db 1738 AACACAGCAAAATGCTAAAAAGAGCAATGTGCTTTATGATGATCCTCAAAAGCTTTGGAAG 1797
QY 1821 CAAGTCTGTTATGTGCAAGTCCCAACAGAGTTTGTATGGATGATAGAAATGATCGATAT 1880
Db 1798 AAAACATGTTACGTTACGTTCCCTCAACGGTTTGTATGATGATTTGACATTGACGATCGATAT 1857
QY 1881 GCAACAGGAACACTGTCTTTTTTATATTAACCTTGAAGGGCTTGAAGGCTTCAAGGA 1940
Db 1858 GCTAACCGTAACATTTCTTCTCGATATCAACTTGAAGGGTTGGACGCAATTCAGGGC 1917
QY 1941 CCAGTTTATGTGGAACTGGTTGTTTTCACAGACAGCTATCTATGTTTATGAGGCC 2000
Db 1918 CCAGTTTATGTGGTACCGGTTGTTTTCACAGACAGCGCTATACGGGTATGATCCA 1977
QY 2001 CCAATTAAAGGCGAAGAA-----GCCAGGTTTCTTGCCATCCTATGTGGGGGCAAG 2051
Db 1978 GTCTTAAACAGAGGAGGATTTGGAACCGAATATCATCATCAAGAGCTGTTGCGGCTCGAGG 2037
QY 2052 AAGAGGCAACCAAGTCAAGAAAAAGGAGCTCAGATAGAAAAAGTGCNAACGATGTG 2111
Db 2038 AAAAAGGTAAGGTGGCAACAAGAGTACATTGACAAAAACAGACGACATAAACGGAACC 2097
QY 2112 GACAGTCTGTTCCAGTATTCATCTCGAAGACATAGAGAGGGTGTGTAAGGTGCTGGG 2171
Db 2098 GAATCAACCGCGCCCATTTTCAATATGGAAGATATGAAAGAGGCAATTTGAAGT- - - - - 2151
QY 2172 TTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTTAGAAGAGAGATTTGGCCAG 2231
Db 2152 TACGATGACGAGAGATCTTTTCTCATGGCACA- - - GAGTTACGAAAGCGGTTCCGTTCAA 2208
QY 2232 TCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGGTGTTCCTCAGTCTCCACT 2291
Db 2209 TCCCTGTCTTATTGCTGCCAGCTTCATGGAACAAGCGGCCCTTCTCCTTCCACAAC 2268
QY 2292 CCAGNATCTCTTTGAAGAAAGCTATCCATGTCATAAGTGTGTCGTATGAGGACAAGTCT 2351
Db 2269 TCTGCAACCTCTTTGAAAGAAAGCAATCCATGTTATAGCTGTGGGTACGAGGACAAGACT 2328
QY 2352 GAATGGGGAACCTGAGATTGGTATGATGATCTGTACAGAAAGATATTTCTTACTGGA 2411
Db 2329 GAATGGGCAAGAGATTGGATGGATATATGGATCTGTACGGAAGATATCTTACCCGGG 2388
QY 2412 TTCAGATGTCACGCAAGAGGCTGGCGTTTCACTTATTGTCATGCCCAAGCGCCAGCTTTC 2471
Db 2389 TTCAGATGTCATACGAGAGGATGGATTCAATCTACTATGTTTATCCACTCACCTCCATT 2448
QY 2472 AAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTGCGGTGGGCTCTC 2531
Db 2449 AAGGATCTGCACCAATTAATCTTTTCGGAGCTGTTTGAACCAAGGCTTTCGATGGGCTCTC 2508
QY 2532 GGTTCGTGTGAATTTCTTTTTCAGCGGCAATTCCTTATGCTATGCTATGCTACGAGGAGGCGC 2591
Db 2509 GGATCGATTGAGATTTCTTTGAGTAGACATTTGCCCATTTGGTATGGCTACAGCGGTAGA 2568
QY 2592 CTCAGTTTCTTGGAGAGATTCCGTTTACATCAACACCAACCATTTTACCCACTAACCTCTCTC 2651
Db 2569 CTCAGTTTCTTGGAGAGATTGCTTATATCAATACTATGTTTATCCACTCACCTCCATT 2628
QY 2652 CCGCTTCTAGTCTATGTTATATGCTGCTATCTGCTGCTACTGGAAGTTCATCATG 2711
Db 2629 CTTTACTTGTCTATTTGCAACCTTCTCTGCTATCTGCTTACTCACCGGAAAGTTCATCGTT 2688
QY 2712 CCAGAGATTAGCAACTTGGCCAGTATCTGGGTTCTTCTTCTTCTTCTTCTTCTCGCC 2771

Db 1394 GTCTCCTGCTATGTTTCTGATGTTGCTGCAATGCTTACGTTTGAAGCATTTGCTGAA 1453
QY 1284 ACTTCAGACTTTGCTAGGAATGGGTACCAATTTGTGAAGAAGTATGACATTTGAACCCAGA 1343
Db 1454 ACATCTGAATTTGCAAGAAATGGGTCTCTTCAGCAAAAGTTTAATATCGAGCCTCGT 1513
QY 1344 GCTCCCGAGTTTACTTTTCCAGAAAATTTGATTACCTGAAAGCAAAAGTCCAGCCTTCA 1403
Db 1514 GCTCCTGAGTGTACTTCCACAGAGAGATAGACTACTGAAAGCAAGGTTGCTGTTCA 1573
QY 1404 TTTGTTTAAAGACCCCGGGCCATGAAGAGAGAATATGAAGAATTTAAATACAGGATAAAT 1463
Db 1574 TTTTGTAGGAGAGAGGGCGATGAAGAGAAATACAGAGAAATCAAGGTAAGGATCAAT 1633
QY 1464 GCCCTAGTTTCTAAGGCATTTAAAGTCCCGAGGAGGATGATCATGCAAGATGCACA 1523
Db 1634 GCCTTGTTGCAAAAGCCCAAGAGTTCTTGAGGAGAGGATGGACAATGCAAGATGGAAGC 1693
QY 1524 CCATGGCCAGGAACAAATACAGGGATCAATCCTGGAATGATTCAGGTTTTTCCCTTGGTCA 1583
Db 1694 CCCTGGCCTGGAACAAACGTACGCGATCATCTCTGGAATGATTCAGGTATTTCTTGCCAA 1753
QY 1584 AGTGGTGGCCTTGATACTAGAGTAAATGAGTCCCGGTTAGTTTATGTGTCCTGTA 1643
Db 1754 AGTGGCGGTGCTGATGTGGAAGAAATGAGTTGCCCTGGCTTTATGTCTCGAGAGAA 1813
QY 1644 AAGCGTCTCTGGGTTCCAGCACCAACAGAGGCTGGTGCCATGAATGCCCTTGTCTGTC 1703
Db 1814 AAGAGCCAGGTTATAACCATCACAGAGAGCTGGTGCCATGAATGCATGTCCTGTC 1873
QY 1704 TCAGCTGCTCTTACTAATGACATACATGTTGAAATCTTGAAATTTGATGTGATCACTACATCAAC 1763
Db 1874 TCTGCTGCTTATCAAAATGCTGCATACCTATTGAACCTTGGACTGTGATCACTACATCAAC 1933
QY 1764 AACAGAAAGCTGTCCGAGAGCTATGTGCTTCTTATGGATCCAAACCTAGTCCGCA 1823
Db 1934 AATAGCAAGGCCATAAAGAGGCTATGTGTTTCATGATGATCCTTTGGTGGGGAAGAA 1993
QY 1824 GTCTGTTATGTGAGTTTCCACAAAAGGTTTGATGGGATTTGATAGGAATGATCATATGCA 1883
Db 1994 GTGTCATGTACAGTCCCTCAGAGGTTTGTGATTTGATTTGACAAAATGATCATAGCT 2053
QY 1884 AACAGGAACACTGTCTTTTGTATTAATCTTGAAGGGCTTGACGGCATCTAAGGACCA 1943
Db 2054 AACAGGAACGTTGTCTTTTGTACATCAACATGAAAGGTTTGGACGGTATTTCAAGGACCC 2113
QY 1944 GTTTATCTGGAACTGCTGTTTCAACAGAACAGCTATCTATGTTATGAGCCCCCA 2003
Db 2114 ATTTATGTGGTACTGATGTGTTTCAGACGCGAGCTGTATGTTATGATGCTCCT 2173
QY 2004 ATTAAGCGCAAGACCA-----GGTTTCTTTGGCATCA 2036
Db 2174 AAAACGAAGAAGCCACCATCAAGAACTTGCAACTGCTGGCCCAAGTGGTGCCTCTTTCG 2233
QY 2037 CTATGTGGGGCAAGAGAGGCAAGCAAGTCAAGAAAGGAGCTCAGATAGAAAAA- 2095
Db 2234 TGCTGACAGAGAACAGAAATAAAGAGAGACTACAAAACCAAGACGGAAGAGAGAAA 2293
QY 2096 -----GTCGAACAGCATGTGGACAGTTCTGTTCAGTATTTCAATCTCGAAGACATAGAG 2150
Db 2294 AGATTATTTTCAAGAAAGCAGAAACCCATCTCTGCAATATGCTTTGGGTGAATTTGAT 2353
QY 2151 GAGGGTTTGAAGGTGCTGGGTTTGTATGATGAAATACAGTTCTCATGTCTCAAAATGAGC 2210
Db 2354 GAAGGTGCTCCAGGTGC-----TGATATCGAAGAGCCGGAATCGTAATCAACAGAAA 2407
QY 2211 TTAGAGAGAGATTTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGTT 2270
Db 2408 CTAGAGAGAAATTTGGGCAGTCTTCTGTTTTTGTGCGCATCAACACTTCTTGAGAACGGA 2467
QY 2271 GGTGTTTCTCAGTCTCCACTCCAGAAATCTCTTTTCAAGAGAGCTATCCATGCTCAATGCT 2330
Db 2468 GGGACCTTGAAAGCGCAAGTCCAGGTTCTTCTTGAGGAAGCTATACATGTTATCAGC 2527

QY 2331 TGTGCGTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGGATCTATGGATCTGTGTC 2390
Db 2528 TCGCGCTACGAAGACAAAGACCGACTGGGAAAAGAGATTTGGCTTATTTACGATCGATC 2587
QY 2391 ACAGAAGATATTCTTACTGGATTCAAAGATGCACGCAAGAGGCTGGCGTTCACTATTGTC 2450
Db 2588 ACAGAGGATATCTTGACTGGATTTAAGATGCATGCCATGGCTGGCGGTCTATTACTGC 2647
QY 2451 ATGCCAAAGCCGACGTTTCAAGGGATCTCCCCCATCAATCTTTAGATGCTCTGAAC 2510
Db 2648 ATCCGAAAGCGCCTGCATTTCAAAGGTTCTCGCCTCTGAAACCTTTCCGACCGCTCTCAC 2707
QY 2511 CAAAGTCTGCGTGGGCTCTCGGTTCTGTTGAAATCTTTTACGCGGCAATGGCCCTTA 2570
Db 2708 CAGGTCTTCTGCTGGGCCCTTGGGTCCGTCGAAATTTCTTTCAGCAAGCACTGCCACTT 2767
QY 2571 TGGTATGCTACGAGGGCGCTCAAGTCTCTGGAGAGATTCGCTTATACATCAACACACC 2630
Db 2768 TGGTACGATACGCGGGGGCTAAATTTCTGGAAGGTTTCTTATATCAACTCCATC 2827
QY 2631 ATTTACCCACTAACTCTCTCCGCTTCTAGTCTATTGTATATTGCTGCTATCTGTCTG 2690
Db 2828 GTTTATCCTGGAGCTCCATTTCTCTGCTTACTGTACCTTGGCTGCGCATCTGCTG 2887
QY 2691 CTCCTGAAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCAATTCG 2750
Db 2888 CTCAGGGGAAGTTTATCACACAGAGCTTACCAATGTCCGCCATCTGTTTCATGGCA 2947
QY 2751 CTCCTCTTTCAATTTTCCGCACTGGTATCTTGAATGAGGTGAGTGGTGGGCAAT 2810
Db 2948 CTTTTCATCTGCATCTCGCTGACCGGATCTCTGGAATAGGTGGAGTGGCGTGGCCATC 3007
QY 2811 GACGAGTGGTGGAGAAAGAACAGTCTCGGGTCAATGGAGGTATCTGACATCTGTTT 2870
Db 3008 GACGACTGGTGGAGAACGACAGTCTTGGGTCTATCGGAGCGTTTGGCGCATCTGTTTC 3067
QY 2871 GCGCTCTTTCAGGGTCTTCTGAAGTGTCTGCCGTATCCGACACAACTTTCACCTGTCTCAC 2930
Db 3068 GCGGTGTTCCAGGGCCTCTGAAGGTGTTCCGCGCATCGACAGAGCTTCCACGCTGACG 3127
QY 2931 TCAAAGGCTAATGAGCAAGAGGCGACTTCTGCTGAGCTCTACATGTTCAAGTGACGACG 2990
Db 3128 TCGAAGGCCGGGAGCGACGA---GGAGTTCTCGGAGCTGTACAGGTTTCAAGTGGACCA 3184
QY 2991 CTTCTCATCTCCGACGACCAATTTGATCATTAACATGGTGGTGTGCTGCTGGCGCAC 3050
Db 3185 CTGCTGATACCCCGACCAACCGCTCCTCTGCTGAACTTCATCGGGGTGGTGGCGGGATC 3244
QY 3051 TCCTACGCCATCAACAGTGGTTACCAATCATGGGGCGCTCTTTTGGGAAAGCTCTTCTTT 3110
Db 3245 TCGAAGCGGATCAACAACGGGTACGAGTCTGGGGCCCCCTGTTCGGAAGCTCTTCTTC 3304
QY 3111 GCCTTCTGGGTGATTTGTTACCTTATACCCATTTCTCAAGGTCATTATGGGAGGCAAAAC 3170
Db 3305 GCCTTCTGGGTGATGCTGCCACTGTACCCGTTCTCAAGGGTCTGGTGGGAGGAGAAC 3364
QY 3171 CGCACCCGAGATTGTCATGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3230
Db 3365 AGGACCCGAGATGCTGATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3424
QY 3231 TGGGTCTGTTGATGCCATTCACCTACCGCTCTGCTGGCCCCCAATATCCAAACCTGTGGC 3290
Db 3425 TGGGTCCGCTGACCGGTTCTCCGCAAGAGCAACGCGCGCTCTCTGGAGGAGTGGGC 3484
QY 3291 ATCAACTCTAGGAAAGTGGAG 3313
Db 3485 CTGGACTGCAACTGAAGTGGGG 3507

RESULT 5
US-09-900-237-3
; Sequence 3, Application US/09900237

Patent No. US20020120124A1
GENERAL INFORMATION:
APPLICANT: Allen, Stephen
TITLE OF INVENTION: Plant Cellulose Syntheses
FILE REFERENCE: B01170 US CIP
CURRENT APPLICATION NUMBER: US/09/900,237
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: 60/092,844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US99/15871
PRIOR FILING DATE: 1999-07-13
PRIOR APPLICATION NUMBER: 09/720383
PRIOR FILING DATE: 2000-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Office 97
SEQ ID NO 3
LENGTH: 3776
TYPE: DNA
ORGANISM: Zea mays
US-09-900-237-3

Query Match 36.4%; Score 1320.4; DB 10; Length 3776;
Best Local Similarity 64.5%; Pred. No. 0;
Matches 2135; Conservative 0; Mismatches 1071; Indels 102; Gaps 7;

QY 47 GCTCTGCGCGCCATGGACGGGACGCGGAGCCCTGAAGTCGGGAGGACGCGGGCGG 106
DB 226 GCTCTGCTGATCGCGCGGATGGGAGCCAGGCGGCCAGCCATGGACCGGGAACGG 285

QY 107 GGACGTTGTCAGATCTGCGCCGACGCGCTGGGCACACAGTTGGAGCGGACGCTCTTAC 166
DB 286 CCAGGTTGTCAGATTTGGCGGACGACGCTGGGCGGCAACCGCGGAGCGGCGTTCGT 345

QY 167 CGCTCGGAGCTGCGGCTTCCGGTCTCCGCTGCTCCGCTGCTAGCAGCAGCGCAGGA 226
DB 346 GGCCTGCAAGAGTGGCGCTTCCCGATCTCCCGGACCTGCTACGAGTACGAGCGCGCGA 405

QY 227 GGCACCCAGGCTGCTCCAGTGCAGACAGCAAGTCAAGCGCCACGAGGAGCGCCAGC 286
DB 406 GGCACGACAGTGCCTCCCGAGTGCAGACCGCTTCAAGCGCTCAAGGGGTGCGGCG 465

QY 287 GATCGGGGGAGGAGGACACACTGATGCGGATGATGCTGAGTGTGACTTCACTACCC 346
DB 466 CGTGGCGGGGACGAGGAGGAGCGGCGTCCGACGCTGGAGAACGAGTTCAAC----- 520

QY 347 TGCATCTGGCACTGAGGACAGAGCAGAGATTGCTGACAGATGCGCAGCTGGCGCAT 406
DB 521 ----TGGAGGCAAGCACACTCCAGTACCTCGCGAGTCCATGCTCCACGCCACAT 576

QY 407 GAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCT 466
DB 577 GAGCTACGGCGGCGGCGGACCTCGACGCGGTGCCGACCCATTCCACCCCATCCCAA 636

QY 467 CTCGAAGTATGACAGTGGAGATCCCTAGGGATAGCTCCCTTCAGTCCACCAACGCCA 526
DB 637 TGTTCCCTCCTCACCACAGGACAGATGGTTCGATGACATCCCGCGGACGACGCGCCT 696

QY 527 GATGTACAGGAATCCCTGGAGCTTCGCTGATCATCATGATGTCCTCCCTACCGGGAA 586
DB 697 TGTCCTCTCGTGGTGGCGGGGGAAGAGGATTACCCCTCT----CCCGTACGGGA 753

QY 587 CATCAGACAGCTGCTCCGTTCCCTATGTAATCATACCAAAATCCGTCAAGGGAGTT 646
DB 754 TCCCAACCTCTGTGTCAACCGAGGTCTATGGACCCCTTCCAAAGATCTCCCGCATATGG 813

QY 647 CTCGGCAGTATTTGGGAATCTCCCTGGAAGAGAGATTGATGGCTGGAATGAAGCA 706
DB 814 CTAC-----GGGAGCGTAGCATGGAAGGAGGAGGATGGAGCTGGAAGCAGAAGCA 864

QY 707 GGACAAGGGTGCATTTCCCATGACTAATGGGACAAGCATTTGCTCCCTCTGAAAGTCCGGC 766
DB 865 GGAGAGATGCACACGAGGACGATGGCGGGCGGATGGTGTATGA----- 915

QY 767 AGCTACTGACATCGATGCACTCT 826
DB 916 -----TGCAGATCTACCACTAATGGATGA 939

QY 827 AACTCGCCAGCCCTCTATCTAG 886
DB 940 AGCTAGACAGCCATTTGCCAG 999

QY 887 AATGGTCAATTTGCTCGGGTTC 946
DB 1000 GATGATATAATAATTCGGCT 1059

QY 947 TCCTGTGCGTAATCATACACCA 1006
DB 1060 TCCGGTGCCTGATGCAATTCCT 1119

QY 1007 TTTATCCTGGATACTGGATCAG 1066
DB 1120 CATGCTCTGGATTTCTTGACCAG 1179

QY 1067 TGATAGACTGGCTTTAAGGTAT 1126
DB 1180 TGACCGGCTGAGTTTAAGGTT 1239

QY 1127 CATATTTGTGACGACGACGAC 1186
DB 1240 TTTCTTTGTGACGACGCTTGT 1299

QY 1187 GCTATCCATTTCTGCTGTTGAT 1246
DB 1300 TCTATCTATCTTTCCGGTGGAT 1359

QY 1247 CGGAGCTTCAATGCTGACTTTT 1306
DB 1360 TGGTGCCTGACGCTGCTGACATT 1419

QY 1307 GGTACCATTTGTGAAGAATGAT 1366
DB 1420 GGTTCCTTTGCAAAAAGATAT 1479

QY 1367 GAAATTTGATTACCTGAAAGAC 1426
DB 1480 GAAGATAGACTCTGAAAGAC 1539

QY 1427 GAAGAGAAATTAAGAATTT 1486
DB 1540 GAAGAGAGATATGAGGAATTC 1599

QY 1487 AGTCCCGGAGAGGATGGATG 1546
DB 1600 GGTTCCTGAGAGGATGGAC 1659

QY 1547 GGATCATCTCTGGAATGATTCA 1606
DB 1660 TGATCATCTCTGGAATGATTCA 1719

QY 1607 TAATGAGCTCCCCGTTTAGTT 1666
DB 1720 AAATGAGCTGCTCGATTGGTT 1779

QY 1667 CAAGAAGGCTGGTCCCATGAT 1726
DB 1780 CAAGAAGGCTGGTCTATGAA 1839

QY 1727 ATACATCTGGAATCTTGATTG 1786
DB 1840 TTATTTGCTGAATTTGATTC 1899

QY 1787 TATGCTTCTTAATGATTCCT 1846
DB 1900 AATGCTTTTATGATGATTCCT 1959

QY 1847 AAGTTTTCATGGGATTCATAG 1906

Db 1960 AGATTGATGGATGATCGCCATGATCGATATGCTACAGAAATGTTGCTTTTCGA 2019
QY 1907 TATTAATCTTGAGGGGCTTTGACGGCATTCAGGACAGTTTATGTTGGAACTGGTTGTGT 1966
Db 2020 TATCAACATGAAAGTTTGGATGTTATCCAGGGCCAAATTTATGTTGGGTACTGGATGTT 2079
QY 1967 TTTCAACAGAACAGTATCTATGTTGTTATGAGCCCCCAATTAAGGGGAG- 2015
Db 2080 CTTTCAAGAGCAGCATATATATGTTGCTATGCTTCCCAAAACAAAGAACCCACCATCAAG 2139
QY 2016 -----AAGCCAGTTTCTTGGCATCACTATGTTGGGGCAAGAA 2056
Db 2140 AACTTGCACACTGCTGCCAAAGTGGTGCAATTTGCTGTTGCTGTTTGGTAACAGGAAGAC 2199
QY 2057 GGAACCAAGTCAAGAAAGAGGAGTTCAGATAAGAAAAAGTCAAGCAAGCATGTGACAG 2116
Db 2200 CAAGAAGAGAACCAAGACCTTAACCTAAATTTGAGAAGATAAAGAAACTTTTAAAGAA 2259
QY 2117 TTTCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTGTTGAAGGTGCT- 2173
Db 2260 AAGGAAAAATCAAGCCCTCGCATATGCTTGTGAAATTTGATGAAGCCGCTCCAGGAGC 2319
QY 2174 TGATGATGAGAAATCAGTTCTCATGTCCTCAATGAGCTTAGAGAGAGATTTGGCCAGTC 2233
Db 2320 TGAATGAAAGGCTAGTATTGTAATCAACAGAGTTGGAAGAAATTTGGCCAGTC 2379
QY 2234 AGCAGCATTTGTTGCCCTCCACTCTGATGGAATATGGTGTGTTCTCAGTCTCCACTCC 2293
Db 2380 TTCAGTTTTTGTGTCATCCACACTTCTTGAGAATGTTGGAACCTGAAGAGTGCCAGTCC 2439
QY 2294 AGAATCTCTTTGAAAGAGCTATCCATGTCATAGTTGTGCTATGAGGACAAGTCTGA 2353
Db 2440 AGCTTCTCTTGAAGAGCTATACATGTCATGTTGTGATATGAAGACANAACAGG 2499
QY 2354 ATGGGAACCTAGATGTTGGATCATGATCTGTCACAGAGATATTTCTTACTGGATT 2413
Db 2500 CTGGGAAAGATATTGTTGGATTTATGGATCAGTCACAGAGATATTTCTTACTGGTT 2559
QY 2414 CAAGATCGACGCAAGAGGTGGGGTTCAGTCTATTCATGTCGCCAACAGCCCGCCTTCAA 2473
Db 2560 TAAGATCGCATCGCATGTTGGGGTCAATTTACTGATACCTAAACGGGCGCCCTCAA 2619
QY 2474 GGGATCTGCCCATCAATCTTTTCAGATCGTCTGAACCAAGTCTGCGTGGGCTCTGG 2533
Db 2620 AGGTTCGCGACCTCTCAATCTTTCCGATCGTCTACCAGGTCTTCGGTGGGCTCTGG 2679
QY 2534 TTTCTGTGAAATCTTTTCAGCGCGCATTTGCCCTTTATGGTATGGCTACGGAGGCGCT 2593
Db 2680 TTCAATGAAATTTTCTTCAGCAACCACTGCCCTCTCTGATGGTATGTTGGTGA 2739
QY 2594 CAAGTTCTCGGAGAGATTCGCTTACATCAACACCACCATTTACCCTAACCCTCTCTCC 2653
Db 2740 AAGATCTCTGGAAGGTTTTCGTACATTAACCTCGTATACCTTTGGACATCTATCC 2799
QY 2654 GCTTCTAGTCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2713
Db 2800 GCTCTGGCCATTTGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2859
QY 2714 AGAGATTAGCAACTTTGGCCAGTATCTGTTGTTCAATGTTGCTTCTTCAATTTTCCGCC 2773
Db 2860 AGAGCTTAACAAATGTTGCCAGCTCTGTTTCATGTCATCTTTCATCTGCTATTTTGTCTAC 2919
QY 2774 TGTATCTCTTGAAGATGAGTGGTGTGTCATGAGAGTGGTGGAGGAATGAACA 2833
Db 2920 GAGCATCTTGAATGAGATGGAGTGGTGTAGGCATCGATGCTGTTGAGAAACGAGCA 2979
QY 2834 GTTCTGGGTCATTGGAGGTATCTCTGACATCTGTTTGGCGTCTTTTTCAGGGCTTCTGAA 2893
Db 2980 GTTTTGGTTCATTGGAGGCTGTCTTCATCTCTTTGCTGCTTCCAGGGACTCTCTCA 3039
QY 2894 GTTGTCTGCGGTATCGACACCAACTTCTACTGTCACCTCAAGGCTTAATGACGAGAGG 2953

Db 3040 GGTCAATAGCTGGTGTAGACAGAGCTTCACTGTGACATCCAAGGCGGAGACGACA 3096
QY 2954 CGATTTTGTGAGCTTACATGTTCAAGTGGACGAGCTTCTCATCCCTCCGACGACCAT 3013
Db 3097 GGAGTTTCTCAGAGCTGTACACATTTCAATGGAGACCTTCTGTATACCTCGACAACCT 3156
QY 3014 TTTGATCATTAACATGTTGTTGCTGCTGTCACCTCTCTAGCCATCAACAGTGGTTA 3073
Db 3157 GCTCCTACTGAATTCATTTGAGTGGTGTAGCTGGCATCTCCAATGGCATCAACACGGATA 3216
QY 3074 CCAATATAGGGGGCGCTCTTTTGGGAAGCTTCTTTTGGCTTCTGGGTGATTTGTTCACTT 3133
Db 3217 TGAATCATGGGGCCCTCTTTCGGGAAGCTTCTTTTGGTATTTGGGTGATCTGCCATCT 3276
QY 3134 ATACCCATTCCTCAAGGCTTATGGCAGGCAAAACACACACGACGATTTGTCATCGT 3193
Db 3277 TTACCCGTTCTCAAGGCTGTTGGAGGCAAAACAGCAGCCACACGATTTGTCATTTG 3336
QY 3194 CTGGGCTGCTCCTCCTGCTTCTATCTTCTCTGCTGCTGGTCTGCTGCTGCTGCTGCTGCT 3253
Db 3337 CTGGTCCATCCTCCTGCTTCTGATCTTCTGCTGCTTGGGTCGGGATCGGATCGGCTTCT 3396
QY 3254 TACCCGCTCTGCTGGCCCAATATCCAAACCTGTGGCATCAACTGTAGGAAAGTGGGAG 3313
Db 3397 TGGCAAGGATGATGGTCCCTGTTGGAGGAGTGTGGTCTGATGCAACTAGGAGGTGAG 3456
QY 3314 TTTGTAGA 3321
Db 3457 CACGTGGA 3464

RESULT 6
US-09-900-237-9
; Sequence 9, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Syntheses
; FILE REFERENCE: B11170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 3936
; TYPE: DNA
; ORGANISM: Zea mays
us-09-900-237-9

Query Match 35.1%; Score 1271; DB 10; Length 3936;
Best Local Similarity 63.9%; Pred. No. 0;
Matches 2110; Conservative 0; Mismatches 1080; Indels 111; Gaps 8;
QY 47 GCTCTGCGCGCCCATGAGCGGACGCGGAGCCCTTGAAGTTCGGGGAGGACGCGGGCGCG 106
Db 191 GCTGCTGTCATCCGCGCGGACGCGGATCCCGGGCCGAAGCCCGCGGGAGCAGAAACGG 250
QY 107 GGAGCTGTCCAGATCTCCCGGAGCGGCTTGGGACCACTGTTGACGCGGACGCTTTCAC 166
Db 251 GGAGTGTGCCAGATTTTGGCGGACGAGCTGGGCTTTGCCCGCGGGGACCCCTTTCGT 310
QY 167 GCGCTGCGACATCTCCCGCTTCCCGGCTGTCGCGCCCTGTCTACGAGCAGCAGGCAAGGA 226
Db 311 GCGGTGCAACAGTGGCGCTTCCCGCTGTCGCGGAGCTGTCTACGAATACGAGGCGCGGA 370
QY 227 GGGACCCAGCGCTTCCCTTCCAGTGAAGCACTAAGGCCACAGAGGAGGAGGCCAGC 286

Db 371 GGGACGCGAAGCTGCCCCCAGTGCAGAGACTCGATACAGCGCCTCAAGGGCTGCCAAGC 430
Qy 287 GATCGCGGGAGGAGCGGACACTGATGCCGATGATGGTAGTACTTCAACTACCC 346
Db 431 TGTGACCGGTCAGGAGGAGGAGCGCGTGCATGACCTGGACAACGAGTTCAACTGGGA 490
Qy 347 TGCATCTGCGCACTGAGGACGAGAGAGAGATGCTGCAGAGTGGCGAGCTGGCGCAT 406
Db 491 CGGCATCACTC-----GCAGTCTGTGGCGAGTCCATGCTCTACGGCCACAT 538
Qy 407 GAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCT 466
Db 539 GAGCTACGGCGGTGGAGTG-----CCCTAATGGCGG 572
Qy 467 CTCCAAGTATGACAGTGGAGAGATCCCTAGGGATAGTCCCTTCAGTACCAACAGCA 526
Db 573 CCACAAGCTTTCCAG-----CTCAACCCCAATGTTCCACTCCTCACCACCGGGA 622
Qy 527 GATGTCAGGAGAAATCCCTGGAGCTTCGCCTGATCATCATCATGATGTCCTCCTACGGGAA 586
Db 623 AATGGTGATGACATCCACCGGAGCAGCAGCGCTGGTGCCTTCTTCATGGGTGGTG 682
Qy 587 CATCAGCAGAGCTGCTCGGTTCCCTATGTAATCATATCACCATAATCCGTCAGGGAGTT 646
Db 683 GGGAAAGAGGATACATCCCTTCCCTTATGCGGATCCAGCTTACCTGTGCAACCCAGGTC 742
Qy 647 CTCGGCGAGTATGGGAATGTTGCTGGAAGAGAGAGTGTGATGGCTGGAAATGAAGCA 706
Db 743 TATGGACCCATCCCAAGATCTGTGATATGGTATGGTAGTGTGCTTGGGAAGCAAG 802
Qy 707 GGACAAGGGTCGATTCCTATGACTAATGGGACAAGCATTTGCTCCCTCTCAAGGTCGGGC 766
Db 803 GATGAGAAATGGAGAGAGAGACAAGAGAGATGCACGACGGGAAATGATGGTGGTG 862
Qy 767 AGCTACTGACATCGATCATCTACTGATACACATCGAAGAGCGCTTACTGATGATGA 826
Db 863 TGATGATGGTGACCATGCT-----GATCACCACATATGATGA 901
Qy 827 AACTCGCGAGCCTCTATCTAGAAAGTCCCAATGCTTCCTCCAAAATAATCCCTACAG 886
Db 902 AGCAAGACAACACTGCCAGGAAATTCACATTCACATCAAGCCAGATTAATCCATATAG 961
Qy 887 AATGGTCATTTCTCGGGTGGTTGTTCTAAGCATCTTCTCGCATCTCCTGCATACCGTCTCAAA 946
Db 962 GATGATTATCATATTCTCGGCTGGTTGGTTTGGGGTCTTCTTCCACTACCGAGTGATGA 1021
Qy 947 TCCTTGCGGTAAATGCATACCCACCTGTGGCTTTATCTGTTATATGTGAGATTTGGTTGC 1006
Db 1022 TCCGGTGAATGATGATTTGCTTTGGCTCATATCTGTTATCTGTAATCTGGTTGC 1081
Qy 1007 TTTATCTGGATAGTGGATGTTCCCGAAGTGGTTTCCCAATCAACCGGGAGACCTACCT 1066
Db 1082 CATGCTCTGGATTTGATCAATTCCTCAAAAGTGGTTCCTTATGAGAGAGACTTACCT 1141
Qy 1067 TGATAGACTGCTTTAAGGTATGACGAGAGAGTGAACCTCTCAGTTGGCTGCTGTTGA 1126
Db 1142 AGACCGGCTGCTACTGAGGTTCGACAAGGAGGCCAGCCATCTCAACTGCTCCCAATTGA 1201
Qy 1127 CATATTTCTCAGTACAGTCCACCCCTTGAAGGAGCCACCTATCGTCACTGCCAACACTGT 1186
Db 1202 TTTCTTGTGATGAGTGTGATCCCTTAAGGAACTCCTTTGGTCACACAATACTGT 1261
Qy 1187 GCTATCCATCTTCTGTTGATATCCCGTGGACAAGGCTCTTGTGATGATGATGATGA 1246
Db 1262 TCTATCTATCCTTTCCGGTGGATATTCCTGTTGATAGGTTTCTGCTATGTTTCTGATGA 1321
Qy 1247 CGGAGCTTCAATCTGACTTTTGACGATTTGGCTGAGACTTCAGAGTTGCTAGGAAATG 1306
Db 1322 TGGTGTGCAATGCTACGTTTGAAGCATTTATCTGAAACATCTGAATTTGCAAGAAATG 1381
Qy 1307 GGTACCAATTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTGGCA 1366
Db 1382 GTTCCCTTTCGCAACCGGTACATATTTGAACCTCGCGCTCCAGAGTGGTACTTCCAACA 1441

Qy 1367 GAAATTCATTACCTGGAAGAG----- 1426
Db 1442 GAAGATAGACTACTTGAAGAG----- 1501
Qy 1427 GAAGAGACAATATCAAGAAATTT----- 1486
Db 1502 GAAGAGAGATGATGAGGAATTT----- 1561
Qy 1487 AGTCCCGGAGGAAGATGGATG----- 1546
Db 1562 AGTTCCCTGAAGAAGATGGACA----- 1621
Qy 1547 GGATCATCTCGAATGATTCAG----- 1606
Db 1622 TGATCATCTCGAATGATTCAG----- 1681
Qy 1607 TAATGAGCTCCCGGTTTAGTT----- 1666
Db 1682 AATGAACCTGCCACGATTTGTT----- 1741
Qy 1667 CAAGAGGCTGGTGCATGAAAT----- 1726
Db 1742 TAAGAAAGCTGGTGCATGAAAT----- 1801
Qy 1727 ATACATGTTGAATCTTGATTT----- 1786
Db 1802 ATATTTGTTAACTTGGATTT----- 1861
Qy 1787 TATGTGCTTCCCTAATGGATCC----- 1846
Db 1862 AATGTGTTTATGATGACCCCT----- 1921
Qy 1847 AAGGTTTCATGGGATGATAG----- 1906
Db 1922 AAGATTTGATGGGATGATCC----- 1981
Qy 1907 TATTAACTTTGAGGGCCCTTGAC----- 1966
Db 1982 TATCAACATGAAGGTTTGGAT----- 2041
Qy 1967 TTTCAACAGACACGCTATCAT----- 2021
Db 2042 ATTTAGAAAGCGAGCATTTAT----- 2101
Qy 2022 ----- 2059
Db 2102 GACTTGCAACTGCTGGCCCAAG----- 2161
Qy 2060 AAGCAAGTCAAGAAAGAGGAG----- 2113
Db 2162 AAAGAAGACTACCAAAACCCAA----- 2221
Qy 2114 CAGTTCTCTCTCCAGTATTCAT----- 2173
Db 2222 GAACCAATCCCTGCATATGCT----- 2275
Qy 2174 TGATGATGAGAAATCAGTTCTC----- 2233
Db 2276 TGAGAATGAAAGGCCGATTT----- 2335
Qy 2234 AGCAGCATTTGTCCTCCACT----- 2293
Db 2336 TTCTGTTTTTGTATCATCCACA----- 2395
Qy 2294 AGAATCTCTTTTGAAGAGCT----- 2353
Db 2396 TGTCTCTCTTTTGAAGAGCT----- 2455
Qy 2354 ATGGGAACTGAGATTTGGTTGA----- 2413
Db 2456 CTGGGAAAGAGATTTGCTGGA----- 2515

1616 TGGAGGCTTCTGATCTGAAGAAACCACTTCCTCGCCTCTTGTATGTTTCCAGAGAA 1675
1646 GCGTCTGGGTTCCAGACCAACAGAGAGGCTGGTGCCATGAATGCCCTTGTGTGCTC 1705
1676 AAGGCGCTGGTTTCAACACCAACAGAGAGGCTGGTCCATGAATGCTCTGGTTGGGTATC 1735
1706 AGCTGTCTTACTAATGAGCAATACATGTTCAATCTTGATTTGATCATCACTACATCAACA 1765
1736 TGCTGTCTCAAAATGCTCTTCAITGTTGAATGATGATGATGATGATGATGATGATGATG 1795
1766 CAGCAAGGCTGTCGAGAGAGATGATGCTTCTTAATGATGATGATGATGATGATGATGATG 1825
1796 CAGCAAGGCTGCCGAGAGGCTGCTCTTCTGATGAGCCACAACTGGGAGAGGTT 1855
1826 CTGTTATGTCAGTTCCACAAAGGTTTGATGGATGATGATGATGATGATGATGATGATG 1885
1856 CTGTTATGTCAGTTCCACAAAGGTTTGATGGATGATGATGATGATGATGATGATGATG 1915
1886 CAGCAAGGCTGTCGAGAGAGATGATGCTTCTTAATGATGATGATGATGATGATGATGATG 1945
1916 CAGCAAGGCTGTCGAGAGAGATGATGCTTCTTAATGATGATGATGATGATGATGATGATG 1975
1946 TTATGTTGGAACTGTTGTTGTTTCAACAGAACAGCTATCTATGTTATGAGCCCCCAAT 2005
1976 ATATGTTGGGACTGGATGTTTTCAGGAGCAAGCTTGTATGGCTATAATCCTCCCAA 2035
2006 TAAGCGGAAGAGCCAGGTTTCTGGCATCACTATGTTGGGGGCAAGAGAGGCAAGCAA 2065
2036 GGGTCCAAAGCGTCCAAAAATGGTAAAGCTGTGTTGGTCC----- 2076
2066 GTCAAAGAAAGGAGCTCAGATAAGAAAGTCCGAAACAGCATGTGGACAGTTCTGTTC 2125
2077 --CGTGTGTTGAAGCCCAAGAGATTAAGAGAGAGATGATGCAATGAGAGGCTGC 2134
2126 AGTATTAATCTCGAAGACATAGAGAGGTTGTTGAAGGTCGTGGTTGATGATGAGAA 2185
2135 AAGCCTAAA-----AGGGATGGATGATGACAA 2161
2186 ATCAGTTCTCTCAATGAGCTTAGAAGAGATGTTGGCCAGTCAGAGCATTTGT 2245
2162 AGAGGTGTGATGTCCTCAATGAAATTTGAGAGAAATTTGACAAATCCTCTATTTTGT 2221
2246 TGCCTCCACTGATGGAATATGTTGTTGTTCTCAGTCCCTCACTCCAGAACTCTTTT 2305
2222 GACTTCTACCTTGATGGAAGAGGTTGTTGCTCTCTTCAAGTCCAGCTGCCCTGCT 2281
2306 GAAAGAGCTATCCATGTCATAGTTCTGCTATGAGGACAGCTGATGGGAACTGA 2365
2282 TAAGAGAGCCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2341
2366 GATTGTTGGATCTGATGATCTGCACAGAGATATTTACTGATTCAGATGCAAGCTGCAG 2425
2342 GCTTGTGTTGATCTATGATCTATCAGAGAGATATCTACAGGTTTAAAGTGCATG 2401
2426 AAGAGCTGGGTTGATGATCTATGATGATGATGATGATGATGATGATGATGATGATG 2485
2402 CCGTGGTGGAGGCTCATTTATGTTATGCCAAAGAGAGCTGCAATCAAGGTTACTGCTCC 2461
2486 CATCAATCTTTCAGATCTGTAACCAAGTCTGCTGGTGGGCTCTCGGTTCTGTTGAAT 2545
2462 TATCACTTGTGATGATCTGTCACAGGTTCTTCGTTGGGACCTTGGTTCCATTTGAT 2521
2546 TCTTTTACGCGGATCTCCCTTATGATGCTT---ACGAGGCGGCTCCCAAGTTCTCT 2602
2522 TTTCTTTAGTCAACATTCCTCTATGATGATGATGATGATGATGATGATGATGATGATG 2581
2603 GGAGAGATTCGTTTACATCAACACCATTTACCCTAACCCTCTCCCGCTTCTAGT 2662
2582 TGAGAGATTTTCCCTTATGCAACAACTGCTATCCTATCCATCCCTCTCTAGTTGC 2641
2663 CTATGTTATATGCTGCTATCTGCTCACTGGAAGTTTCATCATGCGCAGAGATTAG 2722

2642 CTACTGTATTTCTCCAGCAGTT 2701
2723 CAACCTTGCCAGTATCTGGTTG 2782
2702 CACCTTTGCTGGTTGTACTTT 2761
2783 TGAGATGAGTGGAGTGGTGT 2842
2762 TGAGTTGAATGGAGTGGAGT 2821
2843 CATTGGAGGTATCTGTCACAT 2902
2822 CATTGGTGGTATCAGCTCA 2881
2903 CGGTATCGACACCACTTCAC 2962
2882 TGGAAATGACACCAATTTCA 2938
2963 TGAGCTCTACATGTTCAAGTG 3022
2939 AGAATTGTACACCTTTAAGTG 2998
3023 TAACATGTTGGTGTCTGTTGC 3082
2999 TAACATGTTGGTGTCTGTTGC 3058
3083 GGGCGGCTCTTTGGGAGCT 3142
3059 GGGACACTCTTTGGAAAGCT 3118
3143 CCTCAAGGCTTATGGGCAG 3202
3119 CTTTAAAGTTGATGGTGC 3178
3203 CCTCTCGCTCTATCTCTCTC 3262
3179 GCTATGGCCCTATTTCTC 3238
3263 CGCTGGCCCAATATCCAAAC 3315
3239 TAAGGGACCTGATACCAAGCT 3291

9
se Synthase and Promoter
/838,539
987
048
sion 4.0
gonucleotide
ore 1215.6; DB 10; Length 3328;
ed. No. 0;
Mismatches 739; Indels 21; Gaps 3;
CCAGGCTCTATCTAGAAAAAGTCCCATTCCTTCC 866

Qy	3024	AACATGGTTGGTCTCGTTTCTGCTGGCACTCTCTACGCCATCAACAGTGGTTACCAATCATGG	3083
Db	2819	AACATGGTTGGTCTCGTTTCTGCGGATTTCTCCGATGCCCTCAACAAGGGTACGAAGCTTGG	2878
Qy	3084	GGCCCGCTCTTTGGGAAGCTCTTCTTGGCTTCTCGGTGATTTGTTCACTTATACCAATTC	3143
Db	2879	GGACCACTCTTTTGGCAAGTGTCTTTTCTTCTGGGTCACTCCCATCTTTATCCATTC	2938
Qy	3144	CTCAAGGCTCTTATGGCAGGAAAACCCACACGACGATTTGTTCATCGTCTGGGCTGTC	3203
Db	2939	CTCAAGGCTCTTATGGAGCGCCAAACAGGACACCAACCATTTGTCTCTTTGGTCAAGT	2998
Qy	3204	CTCCTCGCTTCTATCTCTCTCTTCTGCTGTGGTTCGTGTGTATCCATTCACCTACCCGCTC	3263
Db	2999	TTGTTGGCTTCTCTCTCTCTCTCTTGTGGTTCGGATCAACCCGTTTGTGCAGCACCGCC	3058
Qy	3264	GCTGGGCCAA	3273
Db	3059	GATAGACCA	3068
RESULT 9			
US-09-938-842A-559			
; Sequence 559, Application US/09938842A			
; Patent No. US20020160378A1			
; GENERAL INFORMATION:			
; APPLICANT: Harper, Jeff			
; APPLICANT: Kreps, Joel			
; APPLICANT: Wang, Xun			
; APPLICANT: Zhu, Tong			
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS			
; FILE OF INVENTION: SAME, AND METHODS OF USE			
; FILE REFERENCE: SCRIPT300-3			
; CURRENT APPLICATION NUMBER: US/09/938,842A			
; CURRENT FILING DATE: 2001-08-24			
; PRIOR APPLICATION NUMBER: US 60/227,866			
; PRIOR FILING DATE: 2000-08-24			
; PRIOR APPLICATION NUMBER: US 60/264,647			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/300,111			
; PRIOR FILING DATE: 2001-06-22			
; NUMBER OF SEQ ID NOS: 5379			
; SEQ ID NO 559			
; LENGTH: 3255			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
US-09-938-842A-559			
Query Match 32.9%; Score 1192.6; DB 9; Length 3255;			
Best Local Similarity 68.6%; Pred. No. 0;			
Matches 1677; Conservative 0; Mismatches 754; Indels 12; Gaps			
Qy	817	TGAATGATGAACATCCCGACGCTCTATCTAGAAAAAGTCCCATTTGCTTCTCCCAAAATAA	876
Db	761	TGATGGATGAGGAACAACACCTCTCTCAAGAAAGCTACCTATTCTGTTCAAGACAGATAA	820
Qy	877	ATCCCTACAGAAATGGTCATTTGTTCTGCGGTTGGTTGTTCTAAGCATCTTCTGCACTACC	936
Db	821	ATCGTTACAGGATGTTAAATCTGTGTGCGCTCGCATTTCTTGTCTTTTCTTTTCAATATA	880
Qy	937	GTCTCAAAATCTGTGCGCTAATGATGATACCCCACTGTGCTTTTATCTGTTATATGTGAGA	996
Db	881	GAATTCCTCAATGATGATGATATGTTAATGTTAACGTCAAGTTATATATGCGAGA	940
Qy	997	TTTGGTTTGGCTTTATCCTGGATCTGGATCAGTCTCCCGAAGTGGTTTCCCAATCAACCGGG	1056
Db	941	TATGGTTTGCAGTGTCTTGATTTCTGTATCAATTTCCCAATGGTATCCATATAGAAGT	1000
Qy	1057	AGACCTACCTTGATAGACTGGCTTTTAAGTATGACCGAGAGGTGAACCGTCTCAGTTGG	1116
Db	1001	AAACATACCTCGATAGACTCTCTCTCAGTACGAGAGGAAGAAACCGTCAGGATTAG	1060

Db 2141 GATCTGAAGCAACAATGAAATTTGGAGAAGAAGTTTGACAATCTCCGGTTTCGTTG 2200
Qy 2248 CCTCACTCTGATGGAATATGGTGTCTCCTCAGTCTCCTCACTCCAGAACTCTCTTTTGA 2307
Db 2201 CCTCTGCTGTCTACAGACGGTGGAGTTCCCGGTACGCAAGCCCGCATGTTTGTAA 2260
Qy 2308 AAGAAGTATCCATGTCTAAGTTGTGGCTATGAGGACAAGTCTGAATGGGGAACTGAGA 2367
Db 2261 GAGAAGCCATTCAAGTTATTAGCTGCGGGTACGAAGATAAAACCGATGGGGAAAAGAGA 2320
Qy 2368 TTGTTGATCTATGGATCTGTACAGAAATATCTTACTGGATTCAGATGCACGCRA 2427
Db 2321 TCGGGTGAATTTATGGATCGGTGACTGAAGATATCTTCGACGGGTTCACAGATGCATTGCC 2380
Qy 2428 GAGGCTGGGTTCACTATTTGATGCCCAAGGCCAGCTTTCAGGGATCTGCCGCCA 2487
Db 2381 ATGGATGGAGATCTGTGCTACTGTATGCTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTA 2440
Qy 2488 TCAATCTTTCAGATCGTCTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTGAAATTC 2547
Db 2441 TTAATCTTGTAGATCGTCTCATCAAGTCTTACGTTGGGCTCTGGCTCTGTAGAGATTT 2500
Qy 2548 TTTTCAGCGGCATTCGCCCTTATGGTATGGCTACGAGGGCGGCTCAAGTTCTCTGGAGA 2607
Db 2501 TCTTGACGACATTTCCGATATGGTATGGTTATGGTGGTGTAAATGGTTGGAGA 2560
Qy 2608 GATTGCGTTTACATCAACACACCAATTTACCACATACTCTCTCCCGCTTCTAGTCTATT 2667
Db 2561 GATTCTCTTACATCAACTCTGCTCTATCTCTTGAGCTTTCACATTCCTATTCGCTATT 2620
Qy 2668 GTATATTGCCGTCATCTGCTGCTCACTGGAAGTTTCATCATGCCAGAGATTAGCAACT 2727
Db 2621 GTTCTCTCCCGCGGTTTGTACTCACAGGAAATTCATCGTCCCTCGAGATAAGCAACT 2680
Qy 2728 TGGCAGTATCTGTTGCTGCTCTCTCTTTCATTTTCGCCACATGGTATCTCTGAGA 2787
Db 2681 ACGCAGGTATACTCTTCATGCTCATGTTCATATCCATGACGAGTAAGTGAATCCTCGAAA 2740
Qy 2788 TGAGTGGAGTGGTGTGGCATTTGACGAGTGGTGGAGGATGAACAGTCTCTGGGTCAATG 2847
Db 2741 TGCAATGGGAGGTGTGCGAATCATGATGTTGGTGGAGAAACGAGCAGTTTTGGGTAATCG 2800
Qy 2848 GAGGTATCTGTCACATCTGTTTCCGTCCTTTCAGGGTCTCTCAAGGTGCTTCCCGGTA 2907
Db 2801 GAGGGCCCTCTCGCATCTATTGCTCTGTTTCAAGGTTTGGCTCAAAGTCTAGCCCGAG 2860
Qy 2908 TCGACACCAACTCTCACTCACTCAAGGCTAATGACGAAGAGGCGACTTTTGGTCTGAGC 2967
Db 2861 TTAACACGAATTTACAGTCACTTCAAAAGC---AGCAGACGATGGAGCTTCTCTCAGC 2917
Qy 2968 TCTACATGTTCAAGTGGAGCGGCTTCTCATCCCTCCGACGACCATTTTGATCATTTAACA 3027
Db 2918 TTTACATCTTCAAGTGGACAATTTGTTGATTTCTCCGACAAACACTCTGATCATTTAACA 2977
Qy 3028 TGGTTGGTGTCTGCTGGCACCTCTACGCCATCAACAGTGGTTACCAATCATGGGGGC 3087
Db 2978 TCAATTTGGAGTTATTGTGCGGCTTCTGATGCCATTAGCAATGGCTATCACTCATGGGGAC 3037
Qy 3088 CGCTCTTTGGGAAGCTCTTTTTCCTTCTGGGTGATTTGTTTCACTTTATACCATTCCTCA 3147
Db 3038 CTCTCTTTGGGAGACTTTTCTCTGCTCTTTGGGTCAATTTGTTCAATTTATACCAATTCCTCA 3097
Qy 3148 AGGGTCTTATGGGAGGCAAAACCGCACACCGACGATGTGTCATGCTGTGGGCTGTCTCC 3207
Db 3098 AGGGAATGCTTTGGGAAGCAACAAAATGCGTACGATTTATTGGTGTCTGTTCTATTCTTC 3157
Qy 3208 TCGCTTCTATCTTCTCCCTGCTGCTGGGTTCTGTTGATGCCATT 3250
Db 3158 TAGCTTCGATCTTGACACTCTTTGGGGTCAAGTTAACCCTTT 3200

US-09-938-842A-2530
; Sequence 2530, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938.842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2530
; LENGTH: 3255
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2530

Query Match 32.9%; Score 1191.4; DB 9; Length 3255;
Best Local Similarity 68.8%; Pred. No. 0;
Matches 1687; Conservative 0; Mismatches 746; Indels 18; Gaps 3;

Qy 815 ACTGAATGATGAACTCGCCAGCCCTCTACTAGAAAAGTCCCAATTTGTTCTCCTCAAAAT 874
Db 756 AATGATGATGAGGAGGAGCCCATTTGCTAGGAGATACCAATCAATCAGCAAGAT 815
Qy 875 AAATCCCTACAGAAATGTCATTTGTTCTGGGTTGGTTGTTCTAAGACATCTTCTCCTCACTA 934
Db 816 AAATCCTTACCAGATTTAAATGTCAGTCTTGTGATTCTTGTGCTCTCTCTTCTTCACTA 875
Qy 935 CGCTCTCAAAATCTGTCGGTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTGA 994
Db 876 CGGTATCTTACCCTCGTCAAGATGCATATGCTTTGGGCTTATTTCTGTTTATGTGA 935
Qy 995 GATTTGGTTTGGCTTTATCTCTGGATACGTGTCAGTTCCTCCGAAAGTGGTTTCCAAATCAACCG 1054
Db 936 GATATGTTTGGCTTTTTCATGGTCTTTCATCAGTTCCTCAATGTCACCTTATCGAGCG 995
Qy 1055 GGAGACCTTACCTGTAGACTGGCTTTAAGGTATGACCGAGAGGTGAACCGTCTCAGTT 1114
Db 996 AGAAAGCTACTTGGACCGACTCTCATTAAGATATGAGAAAGAAAGGAAACCGTCCGGACT 1055
Qy 1115 GGCTGCTGTGACATATTTGTACGTACACTCGACCCCTTGAAGAGGACCACTTATCGTCAAC 1174
Db 1056 ATCCCTGTGGATGTATTTGTTAGTACATGGATCCATTTGAAGAGCCCTCCGCTTATTAC 1115
Qy 1175 TGCCAAACACTGTGCTATCCATTTCTGCTGTTGATTATCCCGTGGACAAAGGTCTCTTGCTA 1234
Db 1116 TGCAAAATCTGCTCTCTTATTTCTGCTGTTGATTATCTCTCGATAAGGTTGCTGTTTA 1175
Qy 1235 TGTATCTGATGACGAGCTTCAATGCTGACTTTTGACGCAATTTGGCTGAGACTTCAGAGTT 1294
Db 1176 CGTATCTGATGATGGTGTGCTATGCTTACTTTTCGAAAGTCTTTCTGAGACCGCTGAATTT 1235
Qy 1295 TGTAGGAATGGTACCATTGTGTAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTT 1354
Db 1236 CGCAAGAAATGGGTTCCTTCTGCAAGAAATATGTTATTGAGCCTCGTCTGCCGAATG 1295
Qy 1355 TTACTTTTCCAGAAAATTCATTACCTGAAAGACAAAGTCCAGCCCTTCACTTTGTTTAAAGA 1414
Db 1296 GTATTCTGCCATAFAAAATGGACTACTTGAAGAATAAAGTTTCAATCCGCACTTTGTTAGGGA 1355
Qy 1415 CCGCCGGGCATGAAGAGAAATATGAAGATTTTAAATCAGGATTAATGCCCTAGTTTC 1474
Db 1356 GCGCGGCGCATGAAGAGAGATTATGAAGATTTCAAAGATTTCAAAGTAAAGATCAATGCTTTTAGTAGC 1415

Qy 1475 TAAGGCTTGAAGTCCCGAGGAGGATGATCATCAGATGGACACACCATGGCCAGG 1534
Db 1416 AACAGCACAGAAAGTGGCTGAGGATGGTGGACTATGCAAGCGGTACACCTTGGCCCGG 1475
Qy 1535 AAACAATACCAAGGATCATCTCGAATGATTCAGGTTTCTTGGTTCACAGTGGTGGCT 1594
Db 1476 TAATAGTGTGCGAGATCATCTGGCATGATTCAGGTTCTTGGAGTGCAGGTTGCG 1535
Qy 1595 TGATACGTAGGATTAATGAGTCCCGCGTTTGTAGTTATGTTCTGTCGTAAGAGCGTCTGG 1654
Db 1536 TGATGTGGAACAAACAGAGTGGCTCGATTAGTTTACGTTTCTCGTGAAGAGACCGG 1595
Qy 1655 GTTCCAGCACACAAAGAGCTGTGCCATGATGCTTGTGTCGTTCTCAGCTGCTCT 1714
Db 1596 ATTTGATCACCAATGAAGGCTGAGCTATGAATTCCTGTATACGAGTCTCTGGGGTCT 1655
Qy 1715 TACTAATGGCAATPACATGTTGAATCTTGTATGTCATCACTACATCAACACAGCAAGC 1774
Db 1656 ATCAATGCTCTTACCTCTGAATGCGATTGTCATCACTACATCAACATAGCAAGC 1715
Qy 1775 TGTCGAGAGAGCTATGCTTCTTAATGATGATGATGATGATGATGATGATGATGATGAT 1834
Db 1716 TCTAGAGAGCAATGTTTCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1775
Qy 1835 GCAGTTCCTCAAAAGGTTGATGGGATGATAGGAATGATGATGATGATGATGATGATGAT 1894
Db 1776 TCAGTTCCTCAAAAGGTTGATGGGATGATAGGACGATGATGATGATGATGATGATGAT 1835
Qy 1895 TGCTCTTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1954
Db 1836 TGCTCTTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1895
Qy 1955 AACTGGTGTGTTTCAACAGACAGCTATCTATGTTATGAGCCCGCCCAATTAAGCGCAA 2014
Db 1896 TACAGGTTGTTTTCAGGAGCAAGCGCTTACGGATTGATGACACCGCAAGAGAA 1955
Qy 2015 GAAGCCAGG-----TTTCTTGGCATCACTATGTTGGGGCAAGAGAGCAAG 2062
Db 1956 GGGCCCAAGTAAAGACATGCAATGCTGGCCAAATGCTGCTCTCTATGTTTGGTTCAG 2015
Qy 2063 CAAGTCAAAAGAGGAGCTCAGA---TAAGAAAAGTGAACAAGCATGTGACAGATTC 2119
Db 2016 AAAGAATCGTAAAGCAAGACAGTGGCTGCGGATAGAGAGAGAGATAGGAAGCGTC 2075
Qy 2120 TGTTCCAGTATCAATCTCAAGACATAGAGAGGATGTTGAGGTGCTGGGTTGATGA 2179
Db 2076 AAAGCAGATCCAGGCATTAGAAAATATCGAAGAGGCGCGTCACTTAAAGGTTCTAACGT 2135
Qy 2180 TGAGAAATCAGTTCTCATGCTCAAAATGAGCTTAGAGAAGATTTGGCCAGTCAAGCAGC 2239
Db 2136 AGAAGTCAACCGAGGCAATGCAATGAAGTTGGAGAGAAATTTGGGCATCTCCTGT 2195
Qy 2240 ATTTGTTGCTCCTCCTGATGAATATGTTGGTGTCTCTAGTCTCCACCTCCAGAAATC 2299
Db 2196 ATTTGTTGCTCCTCCTGATGAATATGTTGGTGTCTCTAGTCTCCACCTCCAGAAATC 2255
Qy 2300 TCTTTTCAAGAGAGCTATCAATGCTAAGTTGCTGCTATGAGGACAACTGTAATGGG 2359
Db 2256 TCTGCTTAAAGAGGCAATCAAGCTATGTTGCGGATAGAGATTAAGTAAATGTAATGGG 2315
Qy 2360 AACTGAGATTGGTGGATCTATGATCTGTACAGAAAGATATCTTACTGGATTCAAGAT 2419
Db 2316 AAAAGAGATTGGTGGATCTATGTTCTGTACCGAAGATATCTTACGGGTTTAAAGT 2375
Qy 2420 GCAGGCAAGAGGCTGGGTTTCACTGATGATGATGATGATGATGATGATGATGATGAT 2479
Db 2376 GCATCTTCAATGTTGGAGATCTGTTTATGATACACCAAAAGTAGGCGCTTCAAAAGATC 2435
Qy 2480 TGCGCCCATCAATCTTTTCAAGATCTCTGAACCAAGTCTGCGGTGGCTCTCGTTCTGT 2539
Db 2436 AGCTCCAAATCAATCTTTCGATCGCTCTCCATCAAGTCTTCATGAGGCGCTTGGGTCGT 2495

Qy 2540 TGAATATCTTTTACGCGGCAAT 2599
Db 2496 TGAGATTTTCTTGTAGTAGGCAT 2555
Qy 2600 CTTGAGAGATTCCTTACATTC 2659
Db 2556 CTTGAGGCGTGTCTTACATTC 2615
Qy 2660 AGTCTATGTATATATGCTGCT 2719
Db 2616 GCTTACTGTTCTCTCCCTGCT 2675
Qy 2720 TAGCAACTGGCCAGTATCTGG 2779
Db 2676 TAGCAACTATCGAGTATCTCT 2735
Qy 2780 CTTTGAGATGAGTGGAGTGGT 2839
Db 2736 TCTCGAGATGCAATGGGCAAG 2795
Qy 2840 GGTCAATGGAGGTATCTGCAAG 2899
Db 2796 GGTCAATGGAGGTGTTCTGCG 2855
Qy 2900 TSCGCGTATCGACACCACTTCA 2959
Db 2856 TGCTGGTGTGACACTAACTTCA 2912
Qy 2960 TGCTGAGCTCTACATGTTCAAG 3019
Db 2913 CTCTGAGCTTTACCTCTTCAAT 2972
Qy 3020 CATTAACATGTTGGTGTCTGCT 3079
Db 2973 CATAAAGCTCATGAGTCTATAG 3032
Qy 3080 ATGGGGCGGCTCTTGGGAGG 3139
Db 3033 GTGGGACCGCTTTTCGGAAG 3092
Qy 3140 ATTCCTCAAGGCTTATGGGCA 3199
Db 3093 GTTCTTAAAGGTTTGTGCGG 3152
Qy 3200 TGCTCTCTCTCTCTCTCTCTCT 3250
Db 3153 CATCTCTCTGCGCTCGATTCTT 3203

RESULT 11
US-09-900-237-17
; Sequence 17, Application US/099
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: plant Cell
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-
; PRIOR APPLICATION NUMBER: 60/0
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/7
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 2890
; TYPE: DNA
; ORGANISM: Glycine max
US-09-900-237-17

Syntheses

00, 237

4

15871

Query Match		32.6%	Score 1183.4;	DB 10;	Length 2890;
Best local similarity		70.4%;	Pred. No. 0;		
Matches 1631;		Conservative	0;	Mismatches	671;
				Indels	15;
				Gaps	3;
QY	946	ATCCCTGCGGTAAATGATGATCCCACTGGCGTTTATCTCTATATGATGAGATTTGGTTTG	1005		
DB	12	ACCCAGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	71		
QY	1006	CTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1065		
DB	72	CTGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	131		
QY	1066	TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1125		
DB	132	TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	191		
QY	1126	ACATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1185		
DB	192	ACGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	251		
QY	1186	TGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1245		
DB	252	TTCTATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	311		
QY	1246	ACGAGCTTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1305		
DB	312	ATGGTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG	371		
QY	1306	GGGTACCATTTGTAAGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1365		
DB	372	GGGTTCATTTTGTAAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG	431		
QY	1366	AGAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1425		
DB	432	AGAAATGAGTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	491		
QY	1426	TGAAGAGAGATATGAAGATTTAAATCAGGATTAATGCTGATGATGATGATGATGATGATG	1485		
DB	492	TGAAGAGGATTAAGAGATTTAAAGTGAGATTAACAGTTTGGTGGCAACAGCACAA	551		
QY	1486	AAGTCCCGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1545		
DB	552	AGGTTCGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	611		
QY	1546	GGATCATCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1605		
DB	612	GGATCATCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	671		
QY	1606	GTAATGAGTCCCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1665		
DB	672	GAAATGAGTACCCCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	731		
QY	1666	ACAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1725		
DB	732	ACAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	791		
QY	1726	AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1785		
DB	792	CCTATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	851		
QY	1786	CTATGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1845		
DB	852	CTATGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	911		
QY	1846	AAAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1905		
DB	912	AGCGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	971		
QY	1906	ATATTAACATGAGGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG	1965		
DB	972	ATATTAACATGAGGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG	1031		
QY	1966	TTTTCAACAGAACGCTATCTATGTTATGAGCCCCCAATTAAGGGCAA-----GAAGC	2019		

DB	1032	TTTTCAAGAGGTACGACACTTTATGATATGATGATGATGATGATGATGATGATGATGATG	1091		
QY	2020	CAGCTTTCTTGGCATCACTATGTTGGGGCAAGAAGAGGCAAGCAAGTCAAGAAAAGGA	2079		
DB	1092	AACTTTGTAAGTGTGGCCAAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1151		
QY	2080	GCTCAGATAAGAAAAGTGAAGCAAGCATGTGGACAGTCTGTTCCAGTATTTCAATCTCG	2139		
DB	1152	ATGCCAATAGTAAGAGGAGAAAAGAGGTGAAGCAGACAGTGAAGCATCAAGCAGA	1211		
QY	2140	AAGACATAGAGGGGTGTTGAAG-----GTGCTGGTGTGATGATGAGAAATGAGTTT	2193		
DB	1212	TACATGCACTTGAATAATTTGAGCGGGGAATGAAGCAACCAATGAGAACATATCCA	1271		
QY	2194	TCATGCTCTCAATAGCTTTAGAGAAAGATTTGGCCAGTCAGCAGCATTTGTTGCCCTCA	2253		
DB	1272	ATCTGACTCAAAACAAAGTTGGAGAAGAGTTTGGACAGTCTCCAGTATTTGTAGCCTCA	1331		
QY	2254	CTCTGATGGAATATGTTGGTGTCTCTCAGTCTCTCCACTCCAGAAATCTCTTTTAAAGAG	2313		
DB	1332	CACTTTGGATGATGGTGGAGTTCCACATGGCGTGAAGTCTGTCATCCTTTTAAAGAG	1391		
QY	2314	CTATCCATGTCATAAGTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATGGTT	2373		
DB	1392	CCATCCAGTCTCAGTTGTGGTTATGAACACAAACAGAAATGGGAAAGAGTTGGGT	1451		
QY	2374	GGATCTATGATCTGTCACAGAAATATTTACTGATTCAGATGACGACGAGAGGCT	2433		
DB	1452	GGATATATGTTCTGTGACAGAGGATATCTTGACTGGATTTTAAATGCAATGTCATGGT	1511		
QY	2434	GGCTTCAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2493		
DB	1512	GGCGTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1571		
QY	2494	TTTCAGATGCTCTGAAACCAAGTGTGGGTGGCTCTCGGTCTCTGTTGTAATCTTTTCA	2553		
DB	1572	TTTCAGATGCTCTGCAACCAAGTTCTCGGTGGCTCTGTTGTTGATTTTTCAT	1631		
QY	2554	GCGGCACTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2613		
DB	1632	GCAGACATGCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1691		
QY	2614	CTTACATCAACACACACCACTTACCCACTTAACCTCTCTCCGCTCTCTAGTCTATGATAT	2673		
DB	1692	CCTACATTAACCTGCTGATATCCCTGGACTTCCCTCCCATTCCTGCTGCTGCTGCTG	1751		
QY	2674	TGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2733		
DB	1752	TACCAGCCATATGCTTCTGACTGGAAATTTATCGTACCCGAGATTAGCAACTATGCCA	1811		
QY	2734	GTATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2793		
DB	1812	GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1871		
QY	2794	GGAGTGGTGTGGCATGACAGTGGTGGAGGAATGAACAGTCTCTGGGTCAATGGAGGTA	2853		
DB	1872	GGGGCGGTGTAGCATAGACGACTGGTGGAGGAACGACAGTTTGGGTGATCGGAGGTG	1931		
QY	2854	TCTCTGACATCTGTTTGGCTGCTTTCAGGGTCTTCTGAAGGTCTGCTGCGGTATCGACA	2913		
DB	1932	TTTCTTCCCATCTATTTGGCTTATTTGAGGTTTACTGAAAGGCTTGGCTGGTGTGNACA	1991		
QY	2914	CCAACTTCACTGTCACTCAAAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTCTACA	2973		
DB	1992	CAAACTTCACTGTCACTCAAAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTCTACA	2048		
QY	2974	TGTTCAAGTGGACGACGCTTCTCATCCCTCCGACGACCATTTGATCATTAACATGGTTG	3033		
DB	2049	TATTCAGTGGACATCACTCTTGATCCCTCCAATGACTTACTTATCATGAATTTGCTG	2108		
QY	3034	GTGCTGTTGCTGGACCTCTCTACGCCATCAACAGTGGTTACCAATCATGCGGCGCTCT	3093		

Db 2109 GGGTGGTGTTCGGGATCTCAGATGCCATACAAATGGTTATGACTCATGGGACCTCTCT 2168
Qy 3094 TTGGGAAGCTCTTCTTGGCTTTCGGGTGATGTTCACTATATACCCATTTCTCAAGGGTC 3153
Db 2169 TTGGTAGATTGTTCTTGGCATTTGGGTGATPCTTCCATCTACCCCTTCTTGAAGGGGT 2228
Qy 3154 TTATGGGACGACAAACCCGACACCGAGATGTCATGCTGCTGGCTGCTCCTCCTCGCTT 3213
Db 2229 TGCTTGGAAAACAGATAGAAATGCAACCAATATATGGTTTGGTCAATCCTTCTGGCCT 2288
Qy 3214 CTATCTTCTCTCTGCTGCTGGTTCGTTGATCCATT 3250
Db 2289 CCACTTGGACTCTCATGTGGGTGAGAAATTAACCCGTT 2325

RESULT 12
US-09-900-237-23
; Sequence 237, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 2663
; TYPE: DNA
; ORGANISM: Picramnia pentandra
US-09-900-237-23

Query Match 29.5%; Score 1069.6; DB 10; Length 2663;
Best Local Similarity 69.7%; Pred. No. 5.1e-302;
Matches 1514; Conservative 0; Mismatches 634; Indels 24; Gaps 4;

Qy 1094 AGAAGGTGAACCGCTCAGTTGGCTGTTGACATATTTGTCAGTACAGTCGACCCCTT 1153
Db 11 AGAGGGGAAGCCATCTGAGCTAGCTGGCTTAGACATATTTGTCAGTACGGTGGATCCTAT 70
Qy 1154 GAAGAGGCCACTATCGTCACTGCCACACACTGTGCTATCCATTTCTGCTGTGATATCC 1213
Db 71 GAAAGAACCTCCACTTATCACTGCAAACTACTGTGCTATCCATCTTGCAGTTGATATCC 130
Qy 1214 CGTGACAAGCTCTCTGCTATGATCTGTATGATCGGAGCTTCAATGCTGACTTTTGAGGC 1273
Db 131 AGTTGATAAAGTTACATGCTACGTGCTAGATGCTGGTGGCTGCCATGCTTACTTTTGAGC 190
Qy 1274 ATTGGCTGAGACTTCAGAGTTTGTAGAAATGGGTACCATTTGTGAGAAAGTATGACAT 1333
Db 191 ACTTCTGAAACACTCTGAATTTGCAGGAAATGGTCCCTTTCTGTAAGAAGTTTAGCAT 250
Qy 1334 TGAACCCAGAGCTCCGAGTTTACTTTTGGCCAGAAATTTGATTAACCTGGAAGACAAAGT 1393
Db 251 TGAGCCTCGAGCACCAAGATGGTATTTCTCTCAGAAGATGGACTATTTTGAAGAACAAGT 310
Qy 1394 CCAGCTTCATTGTGTTAAAGCCCGGCCATGAAGAGAAATATGAAGAATTTAAAT 1453
Db 311 ACACCATCATTTGTAGGGAAGACGCTGCTATGAAGAGAAATATGAAGTATTCAAAGT 370
Qy 1454 CAGGATAAATGCCCTAGTTTCTAAAGCATTTGAAAGTCCCGAGGAAGATGGATCATGCA 1513
Db 371 TCGGATAAATGGTTGGTGTGCTAGTGGCAAAAGGTTCCCGAGGATGGTTGGACGATGCA 430
Qy 1514 AGATGGCACACATGGCCAGGAACAATACAGGGATCATCTTGAATGATCAGGTTT 1573

Db 431 GGATGGACTCTCTGGCCTGGAA 490
Qy 1574 CTTTGGTCAAGTGGTGGCCTTT 1633
Db 491 CTTTGGTCACAATGGTTCGGT 550
Qy 1634 GTCTCGTGAAAGGCTCCTGGG 1693
Db 551 TTCTCGTGAGAAGAGACTGGC 610
Qy 1694 TGTTCGTGCTCAGCTGCTCCTT 1753
Db 611 GGTACGGTCTCCCGGTTATC 670
Qy 1754 CTACATCAACACAGCAGAGCT 1813
Db 671 TTACATCAACAATAGCAAGCA 730
Qy 1814 AGTCCCGCAAGTCTGTTATGT 1873
Db 731 GGGGAAGAAATTATGCTATGT 790
Qy 1874 TCGATATGCAAAACAGGACACT 1933
Db 791 TCGATATTCCAACCGGAATGT 850
Qy 1934 TCAAGGACCAAGTTTATGTTGG 1993
Db 851 ACAAGGCTATATATGTTGGT 910
Qy 1994 TGAGCCCCCAATTAAGCGAAG 2041
Db 911 TGATGCCACCACTACTAAGAA 970
Qy 2042 TGGGGGCAAGAAAGCAAGCAAG 2101
Db 971 GTGCTGCTGTTGTGGGTCCAG 2158
Qy 2102 ---CAAGCATGTGGACAGTTCT 2218
Db 1031 GTCTAAANAATAGGAAGCATCA 2278
Qy 2159 TGAAGTGTCTGGGTTTGTATGAT 2338
Db 1091 GGGAGGA-----TTGAACCTC 2398
Qy 2219 GAGATTGGCCAGTCAGCAGCA 2458
Db 1145 GAAGTTTGGACAGTCTCCAGTC 2518
Qy 2279 TCAGTCTCCCACTCCAGAACTC 2578
Db 1205 TCAAGATGCAACTCTTCAGCA 2638
Qy 2339 TGAGGACAAGTCTGAATGGGGA 2693
Db 1265 TGAAGATAAACAAGAAATGGGG 2753
Qy 2399 TATTCTTACTCGATTCAAGATG 2813
Db 1325 TATTCTGACTGGATTCAGATG 2873
Qy 2459 GCGCCCAAGTTTCAAGGATCT 2933
Db 1385 ACGCCCGCAITTAAGGGTCT 2993
Qy 2519 GCGGTGGGCTCTCGGTTCTGT 3053
Db 1445 TCGGTGGGCGCTTGGATCTGT 3113
Qy 2579 CTAGGGGCGGCTCAAGTTG 3173
Db 1505 CTATGGGTGGATTAATGCG 3233

Db 968 GTAACCGAAAGAGAGATCAAAATCTAGCAAGAAAGGCTCAGACAGAAAAAATCTAGCA 1027
Qy 2104 AGCATGTGGACAGTCTGTGTCCAGTATTCAATCTCAAGACATAGAGAGGGGTGTGAAG 2163
Db 1028 AGAATGTGACCAACTGTGCCATCTTTAGTCTGAGATATAGAGAGGGGTGGAAG 1087
Qy 2164 GTGCTGGTGTGATGATGAGAAATCAGTTCCTCATGTCTCAATAGCTTTAGAGAAGAGAT 2223
Db 1088 GTGCTGGATTTGATGATGAGAAATCACTACTATTATGTCACAAAATGAGCCCTCGAGAAAAGT 1147
Qy 2224 TTGGCCAGTCAGCAGCATTTGTTCCTCCACTGTGATGAATATGTGTGTCTCAGT 2283
Db 1148 TTGTCTAGTCTGTCTCTTGTGTCTCTACACTGTGAGAAATGTTGGCGTTCCTCAGT 1207
Qy 2284 CCTCCACTCCAGAACTCTTTTGAAGAAGCTATCCATGTCAATAGTTGTGGCTATGAGG 2343
Db 1208 CTGCAACTCCAGAACTCTTCTTAAAGGAAGCTATTATGTATTATCAGTTCTGTGTACGAGG 1267
Qy 2344 ARAAGTCTGAATGGGNACTGAGAT 2368
Db 1268 ATAAATCAGAAATGGGAAGTGAGGT 1292

RESULT 15

US-09-900-237-1

; Sequence 1, Application US/09900237

; Patent No. US20020120124A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen

; TITLE OF INVENTION: Plant Cellulose Syntheses

; FILE REFERENCE: Bb1170 US CIP

; CURRENT APPLICATION NUMBER: US/09/900,237

; CURRENT FILING DATE: 2001-07-06

; PRIOR APPLICATION NUMBER: 60/092,844

; PRIOR FILING DATE: 1998-07-14

; PRIOR APPLICATION NUMBER: PCT/US99/15871

; PRIOR FILING DATE: 1999-07-13

; PRIOR APPLICATION NUMBER: 09/720383

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1221

; TYPE: DNA

; ORGANISM: Hordeum vulgare

; US-09-900-237-1

Query Match 14.5%; Score 524.8; DB 10; Length 1221;
Best Local Similarity 75.5%; Pred. No. 7.4e-143;
Matches 652; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

Qy 2394 GAAGATATTTCTACTGGATTCAAGATGCAGCGAAGAGGCTGGCGTTCAGTCTATTTCATG 2453
Db 5 GAGGATATTTCTACTGGGTTTAAATGCGAAGAGGTTGGATATCAATCTACTACTGATG 64
Qy 2454 CCCAAGCGCCAGCTTTTCAAGGATCTGCCCATCAATCTTTCAGATGCTCTGAACCAA 2513
Db 65 CCACCAGCACTTTTTCAGAGGTTCTGCGCAATCAATCTCTGACCGTCTCAATCAA 124
Qy 2514 GTGCTGGGTGGGCTCTCGGTCTCTGTGTTGAAATCTTTTCAGCGCGGATGGCCCTTATGG 2573
Db 125 GTTCTCGGTGGGCTCTTGGGCTAGTTGAAATCTGTGTTAGCAGACATTTGCTATCTGG 184
Qy 2574 TATGGCTACGAGGGCGGCTCAAGTTCCTGTGAGAGATTCCGCTACATCAACACCATT 2633
Db 185 TACAATTACGGTGGCGGTTGAAACTTCTGGAGAGGATGGCTTACATCAACACCATTGTT 244
Qy 2634 TACCACCTAACTCTCTCCGCTTTCTAGTCTATTGTATATTGGCTGTCTGTCTGCTC 2693
Db 245 TATCCAAATACATCCCTTCCACTTATGCGCTATTGTGCTTCCGTCTCTCTCTCTCC 304
Qy 2694 ACTGGAAGTTTCATCATGCGCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTGGC 2753

Search completed: February 25, 2003, 22:11:45

Job time : 212 secs

Db 305 ACCAACAAATTTATTCATTTCCGAGATCAGTAACATATGCTGGGATGTTCTTATCTTATG 364
Qy 2754 TTCTTTTCAATTTTCGCACTGGTATCTTTGAGATGAGGTGAGTGGTGTGGCATTTGAC 2813
Db 365 TTTGCCCTCCATCTTTGCCACGGGTATATTGGAGCTGGATGGATGGTGTGCGCATCGAG 424
Qy 2814 GAGTGGTGGAGGAATGAACAGTTCTTGGGTCAATTTGGAGGTATCTCTGCAATCTTGTGGC 2873
Db 425 GACTGGTGGAGAAACGAGCAGTTCTGGGTATTGGTGGCACATCTGCCCATCTTTTCGCA 484
Qy 2874 GTCTTTTCAGGCTCTCTGAAAGGTGCTTGGCGGTATGACACCAACTTCACTGTCACTCA 2933
Db 485 GTGTTCAGGGTCTGCTGAAGGTGTTGGCCGGGATTTGACACCAACTTCAAGGTTCAC 544
Qy 2934 AAGCTAATGACGAAGAGGAGCTTTTGTGAGCTCTACATGTTCATGTTGAGGAGCGCTT 2993
Db 545 AAGCAAAACGACGAGGATGGCGATTTTGTGAGTTATACCTGTTCAAGTGGACCACTC 604
Qy 2994 CTCACTCCCTCCGACGACCATTTTGTATCATTAACATGTTGGTGTGCTGGCACCTCC 3053
Db 605 CTCACTCCCTCCGACCACTCCCTTGTGATTAACTGTTGGGATGGTGGCAGGATATCA 664
Qy 3054 TACGCCATCAACAGTGGTTACCAATCATGGGGCGGCTCTTTGGGAAGCTCTTCTTTGCC 3113
Db 665 TATGCCATCAACAGCGGTACCAGTCTTGGGTCCACTCTTCGGAAGCTCTTCTTCTCA 724
Qy 3114 TTCTGGGTGATTTGTTCATCTTATACCAATTCCTCAAGGGTCTTATGGCAGGCAAAACCG 3173
Db 725 ATCTGGGTGATCTCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGGAAGCAGAACCG 784
Qy 3174 ACACCGAGGATGTGTCATGCTGGGCTGCTCCTCCTCGCTTCTATCTCTCTCTCTCTGG 3233
Db 785 ACGCAACCATCGTCATTTGTTGTCATCTCTAGCTCCATCTCTCCCTCTCTCTCTCTGG 844
Qy 3234 GTTCGTGTTGATTCACCTACCTACC 3257
Db 845 GTGAAGATCGACCCCTTTCATATCC 868

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2003, 16:34:11 ; Search time 44 Seconds
(without alignments)
3270.697 Million cell updates/sec

Title: US-09-900-237-30

perfect score:

Sequence: 1 MDGDADALKSGRHGAGDVCQ.....VDPFTTRLAGPNIQTCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs. 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Minimum	DB seq	Length:
Maximum	DB seq	length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Genesec_101002.*

1:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SID52/cgcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	%			ID	Description
	Score	Match	Length		
1	5423.5	93.9	1079	AAU10496	Corn cellulose syn
2	5415.5	93.7	1077	AAU84108	Amino acid sequenc
3	5415.5	93.7	1077	AAU84114	Amino acid sequenc
4	5255.5	91.0	1043	AAU84119	Amino acid sequenc
5	5143	89.0	1076	AAU84121	DNA encoding a mai
6	4530	78.4	1054	AAU848734	Arabidopsis thalia
7	4530	78.4	1065	AAU848733	Arabidopsis thalia
8	4530	78.4	1065	ABU93356	Herbicideally activ
9	4499	77.9	1065	AAU33819	Arabidopsis cellu
10	4209.5	72.9	959	AAU848735	Arabidopsis thalia

PT cellulose in the pericarp, hardening the kernel for improved handling

PS ability

PS Claim 3; Page 85-87; 88pp; English.

CC The invention relates to isolated nucleic acids encoding two cellulose

CC synthase proteins from corn, Cdp45 (cesa-3) and Cqrae19 (cesa-9).

CC Also disclosed are a recombinant expression cassette comprising the

CC polynucleotide (operably linked to a promoter) a host cell comprising

CC the recombinant expression cassette and a transgenic plant comprising

CC the recombinant expression cassette. The nucleic acid is useful in

CC the improvement of stalk quality for improved stand or silage, and in

CC the increased concentration of cellulose in the pericarp, hardening

CC the kernel, and thus improving its handling ability. The nucleic acids

CC may also be used as probes or amplification primers in the detection,

CC quantification or isolation of gene transcripts, as probes in detecting

CC deficiencies in the level of mRNA, for detecting gene mutations or

CC allelic variants, for monitoring up regulation of expression or changes

CC in enzyme activity in screening assays, for site directed mutagenesis,

CC and in sense or antisense suppression of one or more genes in a host

CC cell, tissue or plant. The polypeptides may be used in assays for

CC enzyme agonists or antagonists, and as immunogen or antigen to

CC obtain antibodies specifically immunoreactive with the protein.

CC The present sequence represents a corn cellulose synthase of the

CC invention.

XX Sequence 1079 AA;

SQ

Query Match 93.9%; Score 5423.5; DB 23; Length 1079;

Best Local Similarity 92.2%; Pred. No. 0;

Matches 996; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

QY 1 MDGDADALSKGRHAGDVCQICADGLGTTLDGDVFTACDVCRFPVCRCYEHKEGTQA 60
DB 1 MEGDADGVKSGRRGGQVCQICGCVGTAGDVFTACDVCRFPVCRCYEHKEGTQA 60
QY 61 CLOCKTKYKRRHGSIPAIRGEEDDDTDADDGSDFNYPASGTEQDKOKIADRMRSWRMNTGG 120
DB 61 CPQCKNKYKRRHGSIPAIRGEEDDDTDADDGSDFNYPASGTEQDKOKIADRMRSWRMNTGG 120
QY 121 SGNVGHKPYDSGEIGLSYDSEIPRGYVPSVTSNSQMSGEIPGASPDHMMSPGTCNIGRR 180
DB 121 SGNVGHKPYDSGEIGLSYDSEIPRGYVPSVTSNSQMSGEIPGASPDHMMSPGTCNIGRR 180
QY 181 APFPYVNHSPNPSRFFSGSIGNVANKERVQGMKQDKGAIPTMTNGTSTIAPSEGRAAYTD 240
DB 181 APFPYVNHSPNPSRFFSGSIGNVANKERVQGMKQDKGAIPTMTNGTSTIAPSEGRAAYTD 240
QY 241 DASTYNNMEDALLNDETQPLSRKVPYASSKINPYRMVIVLRLVLSIFLHYRLTNPVNRN 300
DB 241 DASTYNNMEDALLNDETQPLSRKVPYASSKINPYRMVIVLRLVLSIFLHYRLTNPVNRN 300
QY 301 AYPLWLLSVICBIWALSILWLDQFPKWPPIINRETYDLRLALRYDREGSPSOLAAYDIFVS 360
DB 301 AYPLWLLSVICBIWALSILWLDQFPKWPPIINRETYDLRLALRYDREGSPSOLAAYDIFVS 360
QY 361 TVDPLKEPPIVNTANTVLSILAVDYPYDVYSCYVSDGASMLTFDALAETSEFARKWVPFV 420
DB 361 TVDPMKEPPLVNTANTVLSILAVDYPYDVYSCYVSDGASMLTFDALAETSEFARKWVPFV 420
QY 421 KYDIEPRAPEYFQCKIDYLDKQVPSFVKDRAMKREYEEFKIRINALYSKALKVPPEE 480
DB 421 KYDIEPRAPEYFQCKIDYLDKQVPSFVKDRAMKREYEEFKIRINALYSKALKVPPEE 480
QY 481 GHIMQDGTWPGNNTDRHPGMIQVFLHSGGLDTGEGNLPRLVYVSREKRGFQHHKAG 540
DB 481 GHIMQDGTWPGNNTDRHPGMIQVFLHSGGLDTGEGNLPRLVYVSREKRGFQHHKAG 540
QY 541 AMNALVRYSVAVLTNGQYMLNDCDHYINNSKAVREAMCFLMDPNLGPVCVYQPPQRFDG 600
DB 541 AMNALVRYSVAVLTNGQYMLNDCDHYINNSKAVREAMCFLMDPNLGRSVYQVFFQRFDG 600
QY 601 IDRDNDYANRNTVFDDINLRGLDQGPVYVGTGCFVNRNTAIYGEPPPIKAKKPGFLASL 660

DB 601 IDRDNDYANRNTVFDDINLRGLDQGPVYVGTGCFVNRNTAIYGEPPPIKAKKPGFLASL 660
QY 661 CGGKKASKSKSRSDKKKSNKSYPVNLEDEIEEGVEGAGFDDDEKSVLMSQMSLEK 720
DB 661 CGGKKASKSKK -GSDKKKSNKSYPVNLEDEIEEGVEGAGFDDDEKSVLMSQMSLEK 719
QY 721 RFGSAAFAVASTLMEYGGVPOSLLKEAHIVISCGYEDKSEWTEIGWIYGSVTE 780
DB 720 RFGSAAFAVASTLMEYGGVPOSLLKEAHIVISCGYEDKSEWTEIGWIYGSVTE 779
QY 781 ILTGFKHARGWRSVYCMKPREKAPINLSRDLNQVLRWALGVSVEILFSRHCPLMWY 840
DB 780 ILTGFKHARGWRSVYCMKPREKAPINLSRDLNQVLRWALGVSVEILFSRHCPLMWY 839
QY 841 YGGRKLFLERPAYINTTYPLTLYVLCILPAICLLTGKFIIMPEISNFIWISLFI 900
DB 840 YGGRKLFLERPAYINTTYPLTLYVLCILPAICLLTGKFIIMPEISNFIWISLFI 899
QY 901 SIFATGILEMWSGVGIDEMWVAVIGGISAHLFAVFGGLKVLGIDTNETVTSKA 960
DB 900 SIFATGILEMWSGVGIDEMWVAVIGGISAHLFAVFGGLKVLGIDTNETVTSKA 959
QY 961 NDEEGDFAELYMFXTLLIPILINMGVVVAGTSYAINSGYQSWGPLFGKLFPAFW 1020
DB 960 SDEGDFAELYMFXTLLIPILINMGVVVAGTSYAINSGYQSWGPLFGKLFPAFW 1019
QY 1021 VIVHLYPFLKGLMGQRNRTPTAVLASIFSLWVRVDPPTTRLAGPNIOTCGINC 1080
DB 1020 VIVHLYPFLKGLMGQRNRTPTAVLASIFSLWVRVDPPTTRLAGPNIOTCGINC 1079

RESULT 2
AAY84108
ID AAY84108 standard; Protein 7 AA.

XX AAY84108;

XX 03-JUL-2000 (first entry)

XX Amino acid sequence of a maize cellulose synthase.

XX Maize; cellulose synthase; cell quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker.

XX Zea mays.

XX Key Location/

XX Misc-difference 494 /note= "entry by MGG"

XX WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US1876

XX 17-AUG-1998; 98US-009682

XX (PION-) PIONEER HI-BRED IN

XX Dhugga KS, Helentjaris TC

XX WPI; 2000-224343/19.

XX N-PSDB; AA299494.

XX New genes which encode mal

XX useful for modulating the

XX and to produce transgenic

XX Claim 15; Page 95-97; 119p

XX The present sequence repre

XX a maize cellulose synthase polypeptide.

CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.

XX Sequence 1077 AA;

Query Match 93.7%; Score 5415.5; DB 21; Length 1077;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 997; Conservative 46; Mismatches 34; Indels 3; Gaps 2;

QY 1 MGDADALSGRHGADGVCQICADGLTLDGVTACDVCRFPVCRPCYHERKEGTQA 60
DB 1 MEGDADGKSGRGGQVCGDGVGTAEAGDFAACDGVPCVCRPCYERKDGTA 60
QY 61 CLOCKTKYKRHRGSPAIRGEGDDTDADDGSDPNYPASCTEDQOKIADRMRSWRMTGG 120
DB 61 CPOCKTKYKRHRGSPAIRGEGDDTDAD--SDFNLASGNEQDQOKIADRMRSWRMTGG 118
QY 121 SGNVGHPKYDSGEIGLSKYSDEIPRGYVPSVTNSQMSGEIPGASPDHMMSPGTGNIR 180
DB 119 SGDVGRPKYDSGEIGLTKYDSGEIPRGYIPSVTNSQISGEIPGASPDHMMSPGTGNICKR 178
QY 181 APPYVNHSPNSRSESGSIGNVANKERVYDGNKQDKGAIPTWNGTSTAPSEGRAADTI 240
DB 179 APPYVNHSPNSRSESGSIGNVANKERVYDGNKQDKGAIPTWNGTSTAPSEGRVGDVI 238
QY 241 DASTEYNMEDALINDTROPLSRKVP IASSKINPYRMVILRVLSLFLHYRLTNPNVRN 300
DB 239 DASTDYNMEDALINDTROPLSRKVP LPSRINPNRMVILRVLSLFLHYRLTNPNVRN 298
QY 301 AYLPLWLLSVCEIWFALSILDOFPKWFPINRETYLDRALRYDREGPSQLAAVDIFVS 360
DB 299 AYLPLWLLSVCEIWFALSILDOFPKWFPINRETYLDRALRYDREGPSQLAAVDIFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDYPVDKYSCYVSDDGASMLTFDALAETSEFARKWPFV 420
DB 359 TVDPMKEPPIVANTVLSILAVDYPVDKYSCYVSDDGAAMLTFDALAETSEFARKWPFV 418
QY 421 KKYDIEPRAPEFYFCOKIDYLDKQVOPSFVKDRAMKREYEFKIRINALVSKALKVPEE 480
DB 419 KKNIEPRAPEWYFSQIDYLDKQVHPSFVKDRAMKREYEFKVRVNGVLAKAKQVPEE 478
QY 481 GTWMQDGTWPNGNTRDHFCMQVFLGHSGGLDTEGNELPRLVYVSREKRPGFQHHKAG 540
DB 479 GTWMQDGTWPNGNTXDHFCMQVFLGHSGGLDTEGNELPRLVYVSREKRPGFQHHKAG 538
QY 541 ANMALVRVSAVLNNGOYMLNCLDDHYINNSKAVREAMCFMDPNLGPQVCYVQFPQRFDG 600
DB 539 ANMALVRVSAVLNNGOYMLNCLDDHYINNSKALREAMCFMDPNLGRSVYVQFPQRFDG 598
QY 601 IDNRNDYANRNTVFEDINLRGLDGIQGPYVGTGCVFNETALYGYEPPKAKKPGFLASL 660
DB 599 IDNRNDYANRNTVFEDINLRGLDGIQGPYVGTGCVFNETALYGYEPPKAKKPGFLSSL 658
QY 661 CGGKKKASKSKRSDDKSNKHVDSSVPVFNLEDEEGVAGFDEKSVLMSQMSLEK 720
DB 661 CGGKKKASKSKRSDDKSNKHVDSSVPVFNLEDEEGVAGFDEKSVLMSQMSLEK 720

DB 659 CGGRKKASKSKK -GSDKKKSKQHVDSSVPVFNLEDEEGVAGFDEKSVLMSQMSLEK 717
QY 721 RFQGSAAFAVASTLMYGGVQPSSTPESLLKEAHIVISCGYEDKSEWGTGTEIGWYGSVTD 780
DB 718 RFQGSAAFAVASTLMYGGVQPSSTPESLLKEAHIVISCGYEDKSEWGTGTEIGWYGSVTD 777
QY 781 ILTGFKMHARGWSVYCMKPRPAKGSAPINLSDRNLNOVLRWALGSSVEILFSRHCPLWYG 840
DB 778 ILTGFKMHARGWSVYCMKPRPAKGSAPINLSDRNLNOVLRWALGSSVEILFSRHCPLWYG 837
QY 841 YGGRKLKELERFAYINTTIYPLTSLPLVYICILPAICLLTGKFIPEISNLSASFIALEFL 900
DB 838 YGGRKLKELERFAYINTTIYPLTSLPLVYICILPAICLLTGKFIPEISNLSASFISLFI 897
QY 901 SIFATGILEMRWSVGIDEMWRNEQFWIGGISAHLFPAVFOGLLKVLAGIDTNTFTVTSKA 960
DB 898 SIFATGILEMRWSVGIDEMWRNEQFWIGGISAHLFPAVFOGLLKVLAGIDTNTFTVTSKA 957
QY 961 NDEGDEFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINSGYOSWGPLFGKLFPAFW 1020
DB 958 SDEGDEFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINSGYOSWGPLFGKLFPAFW 1017
QY 1021 VIVHLYPFLKGLMGRQNRPTTIVVWAVLLASIFSLLMWRVDPFTTRELAPNTQTCGNC 1080
DB 1018 VIVHLYPFLKGLMGRQNRPTTIVVWAVLLASIFSLLMWRVDPFTTRELAPNTQTCGNC 1077
RESULT 3
ID AAY84114
XX AAY84114 standard; Protein; 1077 AA.
XX AC AAY84114;
XX DT 03-JUL-2000 (first entry)
XX DE Amino acid sequence of a maize cellulose synthase.
XX KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
XX KW transgenic plant; plant breeding marker.
XX OS Zea mays.
XX FH Key Location/Qualifiers
XX FT Misc-difference 494 /note= "encoded by MGS"
XX PN WO200009706-A2.
XX PD 24-FEB-2000.
XX PF 16-AUG-1999; 99WO-US18760.
XX PR 17-AUG-1998; 98US-0096822.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Dhugga KS, Helentjaris TC, Bowen BA, Wang X;
XX DR WPI; 2000-224343/19.
XX DR N-PSDB; AA299512.
XX PT New genes which encode maize cellulose synthase polypeptides in plants
XX PT useful for modulating the expression of cellulose synthase in plants
XX PS and to produce transgenic plants expressing the novel protein -
XX PS Claim 15; Page 142-144; 119pp; English.
XX CC The present sequence represents a maize cellulose synthase polypeptide.
XX CC The cellulose synthase can be used for the improvement of stalk quality
XX CC for improved stand or silage. It also provides an increased concentration
XX CC of cellulose in the pericarp, hardening the kernel and improving its
XX CC handling ability. The sequences are used to produce transgenic plants
XX CC and seeds expressing the cellulose synthase. The polynucleotide is

used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Query Match 93.7%; Score 5415.5; DB 21; Length 1077;
Best Local Similarity 92.3%; Pred. No. 0;
Matches 997; Conservative 46; Mismatches 34; Indels 3; Gaps 2;

QY 1 MGDADALSGRHGAGDVCQICADGLTGLDGVTAACVCRFPVCRPCYHERKEGTA 60
DB 1 MEGDADGVKSGRRGGQVQICGDDGVTGTAEGDVFACDVCVCRPCYERKDGTA 60
QY 61 CLOCTKYKRRHSGPAIRGEEDDDADDGSDFNYPASGTEQKOKIADRMESWRMTGG 120
DB 61 CPOCTKYKRRHSGPAIRGEEDDDAD--SDFNLAGSNEGDDQKOKIADRMESWRMTGG 118
QY 121 SGNVGHKYDSGEIGLSKYDSGEIPRGVPSVTNSQMSGEIPGASPDHMHMSPTGNISRR 180
DB 119 SGNVGRPKYDSGEIGLTKYDSGEIPRGVPSVTNSQISGEIPGASPDHMHMSPTGNIGKR 178
QY 181 APFPYVNSPNRPFSSGNSGNVANKERVDSGKMKQDKGAIPMTNGTSTAPSEGRAADTI 240
DB 179 APFPYVNSPNRPFSSGNSGNVANKERVDSGKMKQDKGTIPMTNGTSTAPSEGRGVGDI 238
QY 241 DASTYNNMEDALLNDETQPLSRKYPVPIASSKINPRMVIVLRLVLSIFLHYRLNPVRN 300
DB 239 DASTYNNMEDALLNDETQPLSRKYPVPISSRINPRMVIVLRLVLSIFLHYRIINPVN 298
QY 301 AYPLMLLSVICBIWFALSWILDQFPKWPPIINRETYDLRLALRYDREGESQAAVDIFVS 360
DB 299 AYPLMLLSVICBIWFALSWILDQFPKWPPIINRETYDLRLALRYDREGESQAAVDIFVS 358
QY 361 TVDPLKEPPIVTANTVLSLAVDYPVKVSCVSDGASMLTFDALAETSEFARKWVPV 420
DB 359 TVDPMKEPPLVTANTVLSLAVDYPVKVSCVSDGGAAMLTFDALAETSEFARKWVPV 418
QY 421 KKYDIEPRAPEFYFCQKIDYLDKQVPSFKVDRRAMKREYEEFKIRINALVSKALKVPEE 480
DB 419 KKYDIEPRAPEFYFSQKIDYLDKQVPSFKVDRRAMKREYEEFKVRVNGLVAKQVPEE 478
QY 481 GHIMQDGTFPWGNNTRDHPCHTQVFLHSGGLDTGNEPLRLVYVSREKRPQFHHKAG 540
DB 479 GHIMQDGTFPWGNNTXDHGMQVFLHSGGLDTGNEPLRLVYVSREKRPQFHHKAG 538
QY 541 ANNALVRYSAVLTNGOYMLNLDCHYINNSKAVRAMCFMIDPNLGPVYVQVQFQRFDG 600
DB 539 ANNALVRYSAVLTNGOYMLNLDCHYINNSKALRAMCFMIDPNLGRSVYVQVQFQRFDG 598
QY 601 IDNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVNRTAIYGEPPPIKAKKPGFLASL 660
DB 599 IDNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVNRTAIYGEPPPIKAKKPGFLSSL 658
QY 661 CGGKKKASKRSKRSKSKNKHVDSSVPVFNLEDIEEGVEGAGDDEKSVLMSQMSLEK 720
DB 659 CGGRKKASKSKK-GSDKKKSKQKHVDSSVPVFNLEDIEEGVEGAGDDEKSLMSQMSLEK 717
QY 721 RFGQNAFVASTLMBYGGVPSQSTPESLLKEAHIYVSCGYEDKSEWGTGIGWYGSVTD 780
DB 718 RFGQNAFVASTLMBYGGVPSQSATPESLLKEAHIYVSCGYEDKTEWGTGIGWYGSVTD 777

QY 781 ILTGFKMHARGWRSVYVCMKRP
DB 778 ILTGFKMHARGWRSVYVCMKRP
QY 841 YGRLKFLERFAYINTTIYPLT
DB 838 YGRLKFLERFAYINTTIYPLT
QY 901 SIFATGILEMRSGVGIDEMWR
DB 898 SIFATGILEMRSGVGIDEMWR
QY 961 NDEEGDAELYMFKWTTLLIPP
DB 958 SDEGDFAELYMFKWTTLLIPP
QY 1021 VIVHLYPFLKGLMGRNRTPTI
DB 1018 VIVHLYPFLKGLMGRNRTPTI
AA.
RESULT 4
ID AAY84119 standard; Protein;
XX AAY84119;
XX AC AAY84119;
XX DT 03-JUL-2000 (first entry)
XX DE Amino acid sequence of a ma
XX KW Maize; cellulose synthase;
XX KW transgenic plant; plant bre
XX OS Zea mays.
XX FH Key Location/Qu
XX FT Misc-difference 494
XX FT /note= "enc
XX PN WO200009706-A2.
XX PD 24-FEB-2000.
XX PF 16-AUG-1999; 99US-US18760
XX PR 17-AUG-1998; 98US-0096822
XX PA (PION-) PIONEER HI-BRED INT
XX PI Dhugga KS, Helentjaris TG,
XX DR WPI; 2000-224343/19.
XX DR N-PSDB; AAZ99527.
XX PT New genes which encode maiz
XX PT useful for modulating the ex
XX PT and to produce transgenic p
XX PS Claim 15; page 181-183; 119
XX The present sequence repres
CC The cellulose synthase can
CC for improved stand or silag
CC of cellulose in the pericar
CC handling ability. The sequ
CC and seeds expressing the ce
CC used for modulating, prefer
CC in a plant cell. The plants
CC is also used as a probe or
CC isolation of gene transcrip
CC deficiencies in the level o

APINLSDRNLQVLRWALGVSVEILFSRHCPLWYG 840
APINLSDRNLQVLRWALGVSVEILFSRHCPLWYG 837
VYCILPAICLLTGKFIPEISNLFASIFSLFI 900
VYCILPAICLLTGKFIPEISNLFASIFSLFI 897
VYGGISAHLFAVFOGLLKVLAGIDNFTVTSKA 960
VYGGISAHLFAVFOGLLKVLAGIDNFTVTSKA 957
INMVGWVAGTSVAINSGYOSWGPLFGKLFFAFW 1020
INLVGVAGISVAINSGYOSWGPLFGKLFFAFW 1017
VILLASIFSLWVRVDPFTTTLRAGPNITQTCGINC 1080
VILLASIFSLWVRIDPFTTTRVTPGTQTCGINC 1077

cellulose synthase.
quality; stand; silage; cellulose;
marker.
ers
by MGG"
en BA, Wang X;
ulose synthase polypeptides in plants
sion of cellulose synthase in plants
expressing the novel protein -
nglish.
a maize cellulose synthase polypeptide.
ed for the improvement of stalk quality
also provides an increased concentration
rdening the kernel and improving its
are used to produce transgenic plants
se synthase. The polynucleotide is
increasing, the level of the synthase
preferably monocots. The polynucleotide
r in the detection quantitation or
he probes are useful in detecting
A in screenings for desired transgenic

CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for
CC recombinant expression of their encoded polypeptides or for use as
CC immunogens in the preparation and/or screening of antibodies. The
CC proteins can be employed in assays for enzyme agonists or antagonists
CC of enzyme function or for use of immunogens or antigens to obtain
CC antibodies specifically immunoreactive with a protein.
XX
SQ Sequence 1043 AA;

Query Match 91.0%; Score 5255.5; DB 21; Length 1043;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 969; Conservative 42; Mismatches 32; Indels 3; Gaps 2;

QY 1 MGDADALSGRHGAGDVCOICADGLGTTLDGVDVCTACDVCRFPVCRPCYEHERKEGQA 60
DB 1 MEGDADGVKSGRRGGVQCGDVGTTAEGDVFCAACDVCGFPVCRPCYEYERKDGTA 60
QY 61 CLOCKTKYKRHGSIPAIRGEEDGDDADGSDFNYPASCTEDOKQKIADMRMSWRMTGG 120
DB 61 CPOCKTKYKRHGSIPAIRGEEDGDDAD--SDFNVLASGNEQDKQKIADMRMSWRMTGG 118
QY 121 SGNVGHPKYDSGEIGLSKYSDEIPRGYVPSVTNSQMSGEIPGASPDHMHMSPTGNISRR 180
DB 119 SGDVGRPKYDSGEIGLTKYDSIEIPRGYIPSVTNSQISGEIPGASPDHMHMSPTGNIGKR 178
QY 181 APPYYNHSPNPSREFSGISGNVAMKERVYDGMKQDKGAIPMTNGTSTIAPSEGRAANDI 240
DB 179 APPYYNHSPNPSREFSGISGNVAMKERVYDGMKQDKGTIPMTNGTSTIAPSEGRGVGI 238
QY 241 DASTENMEDALLNDETROPLSRKVPYIASKINPRRWIVLRVLVLSIFLHYRLNPNVRN 300
DB 239 DASTDYNMEDALLNDETROPLSRKVPYIASKINPRRWIVLRVLVLSIFLHYRLNPNVRN 298
QY 301 AYPWLILLSVCEIWFALSWILOQPKWFFINRETYLDRLALYDREGPSQLAAVDIFVS 360
DB 299 AYPWLILLSVCEIWFALSWILOQPKWFFINRETYLDRLALYDREGPSQLAAVDIFVS 358
QY 361 TVDPLKEPPIVANTVLSILAVDYPVDKVCYVSDDGASMLTFDALAETSEFARKWPFV 420
DB 359 TVDPMKEPPLVANTVLSILAVDYPVDKVCYVSDDGAAMLTFDALAETSEFARKWPFV 418
QY 421 KYVDIEPRAPPEYFCOKIDYLDKQVQSPVKDRRAKREYEFKIRINALSKALKVPEE 480
DB 419 KYNIEPRAPPEYFSQKIDYLDKQVQSPVKDRRAKREYEFKIRINALSKALKVPEE 478
QY 481 GWIMQDCTPWGNTRDHPGMIOVFLGSHSGGLDTEGNELPRLVYVSREKRPQFOHKKAG 540
DB 479 GWIMQDCTPWGNNTXDHFGMLQVFLGSHSGGLDTEGNELPRLVYVSREKRPQFOHKKAG 538
QY 541 AMNALVRVSAVLNTGYMLNLCDDHYINNSKAVREAMCFMDPNLGPQVCYVQFPQREDG 600
DB 539 AMNALVRVSAVLNTGYMLNLCDDHYINNSKALREAMCFMDPNLGRSVCYVQFPQREDG 598
QY 601 IDRNDRYARNRVFFDINRLGLDGTGQPVYVGTGCVFNRTALYGEYPPITAKKPGFLASL 660
DB 599 IDRNDRYARNRVFFDINRLGLDGTGQPVYVGTGCVFNRTALYGEYPPITAKKPGFLASL 658
QY 661 CGGKKKASKSKRSDDKKSNKHVDSSVPVFNLEDEEGVAGDDEKSVLMSQMSLEK 720
DB 659 CGGKKKASKSKK--GSDKKSKQHVDSVVPVFNLEDEEGVAGDDEKSVLMSQMSLEK 717
QY 721 RFQSAAFVASTLMGYGVQSPSTPESLLKEAIIHISCGYEDKSEWGTGIGWYGSVTD 780
DB 718 RFQSAAFVASTLMGYGVQSPATPESLLKEAIIHISCGYEDKTEWGTGIGWYGSVTD 777
QY 781 ILTGFKMHARGVRSVCMKPRAPFAGSAPINLSDRNLQVLRWALGSEVILFSRHCPLMYG 840
DB 778 ILTGFKMHARGVRSVCMKPRAPFAGSAPINLSDRNLQVLRWALGSEVILFSRHCPLMYG 837

QY 841 YGGRLKFLERAYINTTIYPLTSLPLLYCILPAICLLTGKFIIMPEISNLASIWFIALFL 900
DB 838 YGGRLKFLERAYINTTIYPLTSLPLLYCILPAICLLTGKFIIMPEISNFIISLFI 897
QY 901 SIFATGILEMRSWGVGIDEMWRNEQFWVIGGISAHLFAVFGGLKVLGIDTNTFTVTSKA 960
DB 898 SIFATGILEMRSWGVGIDEMWRNEQFWVIGGISAHLFAVFGGLKVLGIDTNTFTVTSKA 957
QY 961 NDEGDFAEFLYFKWTTLLIPPTILLINMVGVVAGTYSAINSGYQSWGPLFGKLFFAFW 1020
DB 958 SDEGDFAEFLYFKWTTLLIPPTILLINVLGVVAGTYSAINSGYQSWGPLFGKLFFAFW 1017
QY 1021 VIVHLYPELKGIMGRQNRTPTIIVVW 1046
DB 1018 VIVHLYPELKGIMGRQNRTPTIIVVW 1043
RESULT 5
ID AAY84121 standard; Protein; 1076 AA.
XX AAY84121;
XX
DT 03-JUL-2000 (first entry)
XX
DE DNA encoding a maize cellulose synthase.
XX
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;
KW transgenic plant; plant breeding marker.
OS Zea mays.
FH Key Location/Qualifiers
FT Misc-difference 425 /note= "encoded by ANG"
XX WO200009706-A2.
PD 24-FEB-2000.
XX
PF 16-AUG-1999; 99WO-US18760.
PR 17-AUG-1998; 98US-0096822.
XX (PION-) PIONEER HI-BRED INT INC.
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
XX WPI: 2000-224343/19.
XX N-PSDB; AAZ99533.
PT New genes which encode maize cellulose synthase polypeptides in plants
PT useful for modulating the expression of cellulose synthase in plants
PT and to produce transgenic plants expressing the novel protein
XX
PS Claim 15; Page 196-199; 119pp; English.
XX
CC The present sequence represents a maize cellulose synthase polypeptide.
CC The cellulose synthase can be used for the improvement of stalk quality
CC for improved stand or silage. It also provides an increased concentration
CC of cellulose in the pericarp, hardening the kernel and improving its
CC handling ability. The sequences are used to produce transgenic plants
CC and seeds expressing the cellulose synthase. The polynucleotide is
CC used for modulating, preferably increasing, the level of the synthase
CC in a plant cell. The plants are preferably monocots. The polynucleotide
CC is also used as a probe or primer in the detection quantitation or
CC isolation of gene transcripts. The probes are useful in detecting
CC deficiencies in the level of mRNA in screenings for desired transgenic
CC plant, for detecting mutations in the gene, for monitoring upregulation
CC of expression or changes in enzyme activity in screening assays of
CC compounds, for detection of any number of allelic variants of the gene,
CC or for use as molecular markers in plant breeding programs. The
CC isolated nucleic acids of the present invention can also be used for

CC	recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.
XX	
SQ	Sequence 1076 AA;
	Query Match 89.0%; Score 5143; DB 21; Length 1076;
	Best Local Similarity 87.6%; Pred. No. 0;
	Matches 947; Conservative 69; Mismatches 59; Indels 6; Gaps 6;
QY	1 MGDADALSKRGHAGDVCQICADGLGTLTDGDFVACDVCPCPCYEHKEGQTQA 60
Db	1 MDG-GDATNSGRHVAGQVQICGDGVTAGDGLFTACDVCPCPCYEHKEGQTQA 59
QY	61 CLQCKTKYKRRHSGPAIRGEEDTDADDGSDFNYPASGTEDQKQIADRMRSRMNTGG 120
Db	60 CPQCKTKYKRRHSGPPVHGEENEDVADDSDYNYQASGDDQKQIAERMLTWRNTRSG 119
QY	121 SGNVGHPKYDSGEIGLSKYSGEIPRGYPVSVTNSQMSGEIPGASPDHMHMSTGNTISRR 180
Db	120 S-DIGLAKYDSGEIGHKYSGEIPRGYIPSLTHSQISGEIPGASPD-HMSPVGNIGRR 177
QY	181 A-PFPYVNSPNPSRPFSGSIGNVANKERVWGKMKQDKGAIPMTNGTSTAPSEGRAATD 239
Db	178 GHQFPYVNSPNPSRPFSGSLGNVANKERVWGWKMK-DKGAIPMTNGTSTAPSEGRVAD 236
QY	240 IDASTEYNMEDALLNDTQPLSRKVPPIASSKINPYRMVILVRLVLSIFLHVRLNPNVR 299
Db	237 IDASTYNNMEDALLNDTQPLSRKVPPISSRNPNRMVILVRLVLCIFLRYRITHPNV 296
QY	300 NAYPLWLLSVICEIWFALSMDQFPKWPPIINRETYLDRLALRYDREGPSSQLAAYDIFV 359
Db	297 NAYPLWLLSVICEIWFALSMDQFPKWPPIINRETYLDRLALRYDREGPSSQLAPVDIFV 356
QY	360 STVDPLKEPPIYNTANTVLSILAVDYPVDKVCYVSDDGASMLTFDALAETSEFARKWVPF 419
Db	357 STVDPMKEPPIYNTANTVLSILAVDYPVDKVCYVSDDGAAMLTFDALSETSEFARKWVPF 416
QY	420 VKKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRRAKREYEFKIRINALVSKALKVPE 479
Db	417 CKKYNIEPAPWYFAQIDYLDKQVTSFVKERRAKREYEFKIRINCLYAKAKVPE 476
QY	480 EGWIMQDGPWGNNTRDHPGMIVQFLHSGGLDTEGNEPLRVLYVYSREKRPQFQHHKA 539
Db	477 EGWIMQDGPWGNNTRDHPGMTIQVFLHSGGLDVEGNEPLRVLYVYSREKRPQFQHHKA 536
QY	540 GAMNALRVYSVAVLTNGQYMLNDCDHYINNSKAVREAMCFMDPNLGPQVCYVQFPQFED 599
Db	537 GAMNALRVYSVAVLTNGQYMLNDCDHYINNSKALREAMCFMDPNLGRNVYVQFPQFED 596
QY	600 GIDRNDYANRNTVFPDINLRGLDGTGGPVYVGTGCVFNRTAIYGEPPITAKKPKGFLAS 659
Db	597 GIDRNDYANRNTVFPDINLRGLDGTGGPVYVGTGCVFNRTAIYGEPPVKKKPKGFFSS 656
QY	660 LCGGKKKSKSKRSSDKKSNKHVDSSVPVFNLEDIEEGVAGDFDDEKSLVMSQMSLE 719
Db	657 LCGGRKTKSKSK-SSEKKKSHRHADSSVPVFNLEDIEEGSQDDEKSLVMSQMSLE 715
QY	720 KREGQSAAFVASTLMYEGVQPSSTPESLLKEAIIHVISGYEDKSEWGTGEIHWIYGVTE 779
Db	716 KRFQGSVFVASTLMYEGVQPSATPESLLKEAIIHVISGYEDKTDWGTGEIHWIYGVTE 775
QY	780 DILITGFKMHARGWRSYCYMKPKPAFGSAPINLSDRNLQVLRWALGSVELFSRHCPWY 839
Db	776 DILITGFKMHARGWRSYCYMKPKPAFGSAPINLSDRNLQVLRWALGSIELFSRHCPWY 835
QY	840 GYGGRUKFLERFAYINTTITPLTSPLLYCYILPAICLLTGKFIIMPEINSLASINFIALF 899
Db	836 GYGGRUKFLERFAYINTTITPLTSIPLLCYILPAVCLLTGKFIIPKISNLESVWFISLF 895
QY	900 LSIFATGILEMRMSGVIGDEWRNEQFWTIGGISAHLPFAVFOGLLKVLAGIDTFTVTSK 959

Db	896 ISIFATGILEMRMSGVIGDEWR	SWVIGGISAHLPFAVFOGLLKVLAGIDTFTVTSK 955
QY	960 ANDEEGDFAELYMFKWTTLLIF	ITINMVGVAAGTSAINSGYOSWGPFLGKLFPAF 1019
Db	956 ATDEEGDFAELYMFKWTTLLIF	ITINLVGVAGISYAINSGYOSWGPFLGKLFPAF 1015
QY	1020 WVIVHLYPFLKGLMGQRNTRPT	SAVLLASIFSLMWVRVDPFTTTRLAGPNITQCGIN 1079
Db	1016 WVIVHLYPFLKGLMGQRNTRPT	SAVLLASIFSLMWVRIDPFTTTRVTGPDIAKCGIN 1075
QY	1080 C 1080	
Db	1076 C 1076	
RESULT 6		
AAG48734		
ID	AAG48734 standard; Protein:	AA.
XX		
AC	AAG48734;	
DT	18-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein	segment SEQ ID NO: 61574.
XX		
KW	Protein identification; sig	transduction pathway; metabolic pathway;
KW	hybridisation assay; genet	oping; gene expression control; promoter;
XX	termination sequence.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439	
XX		
PR	25-FEB-1999;	99US-0121825
PR	05-MAR-1999;	99US-0123180
PR	09-MAR-1999;	99US-0123548
PR	23-MAR-1999;	99US-0125788
PR	25-MAR-1999;	99US-0126268
PR	29-MAR-1999;	99US-0126785
PR	01-APR-1999;	99US-0127462
PR	06-APR-1999;	99US-0128234
PR	08-APR-1999;	99US-0128714
PR	16-APR-1999;	99US-0129845
PR	19-APR-1999;	99US-0130077
PR	21-APR-1999;	99US-0130449
PR	23-APR-1999;	99US-0130510
PR	28-APR-1999;	99US-0130891
PR	30-APR-1999;	99US-0131449
PR	30-APR-1999;	99US-0132048
PR	04-MAY-1999;	99US-0132407
PR	05-MAY-1999;	99US-0132484
PR	06-MAY-1999;	99US-0132485
PR	06-MAY-1999;	99US-0132486
PR	07-MAY-1999;	99US-0132487
PR	07-MAY-1999;	99US-0132863
PR	11-MAY-1999;	99US-0134256
PR	14-MAY-1999;	99US-0134218
PR	14-MAY-1999;	99US-0134219
PR	14-MAY-1999;	99US-0134221
PR	14-MAY-1999;	99US-0134370
PR	18-MAY-1999;	99US-0134768
PR	19-MAY-1999;	99US-0134941
PR	20-MAY-1999;	99US-0135124
PR	21-MAY-1999;	99US-0135353
PR	24-MAY-1999;	99US-0135629
PR	25-MAY-1999;	99US-0136021
PR	27-MAY-1999;	99US-0136392
PR	28-MAY-1999;	99US-0136782

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139483.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143077.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 78.4%; Score 4530; DB 21; Length 1065;
Best local Similarity 78.6%; Pred. No. 0;
Matches 839; Conservative 98; Mismatches 104; Indels 26; Gaps 8;

QY 19 COICADGLGTTLDGDVFTACDVCREFPCRCYEHKEKGTQACLOCKTKYKRRHSGPAIR 78
| | | | | : | : | | | | | | | | | | | | | | | : | : | | | | | : | | | | |
Db 20 CQICSDNVGKTVDCDREFVACDICSFPVCRPCYERKDGNGQPCQCKTKYKRLKGSALP 79
| | | | | : | : | | | | | | | | | | | | | | | : | : | | | | | : | | | | |
QY 79 GEEGDDTDADGS-DFNYPASGTEDQKKTADRMRSWRMNTGGSNVGHKPYDSSGIGUS 137
| : : | | | : | | | | | | | | | | | | | | | : | : | | | | | : | | | | |
Db 80 GDKDEDGLADEGTVFEFNYP-----QKEKISERMLGHLTRGKEEMGEPOYDK----- 127

QY 138 KYDSGEIPRGVYPSVTNSQ-MSGEIPGASPDHMMSPGNGISRRAPPY-VNHSPNPSRE 195
Db 128 -----EVSHNHLPRLTSRQDTSGEFSAASPERLSVSTIAGGKRLPYSSDVNQSPNRRIV 182
QY 196 FSGSTGNVANKRVDGKWKQDKGAIPMTNGTSIAPSEGRATDIDASTEYNMEDALLND 255
Db 183 DPGVLGNVANKRVDGKWKQDKGVPV---STQASE-RGGVDIDASTDILLADEALLND 238
QY 256 ETRQPLSRKVPATSKINPYRMVIVLRLVLSIFLHYRLTNPNVFNAYPLWLLSVICIEWF 315
Db 239 EARQPLSRKVPATSKINPYRMVIVLRLVLSIFLHYRLTNPNVFNAYPLWLLSVICIEWF 298
QY 316 ALSWILDQFPKWPINRETYLDRALRYDREGEPSQLAAVDIFVSTVDPLKEPPIVANT 375
Db 299 ALSWILDQFPKWPINRETYLDRALRYDREGEPSQLAAVDIFVSTVDPLKEPPIVANT 358
QY 376 VLSILAVDYPVKVSCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPRAPERYFC 435
Db 359 VLSILAVDYPVKVSCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPRAPERYFC 418
QY 436 QKIDYLDKVKQPSFVKDRAMKREVEEFKIRINALVSKALKVPEEGIMQDGTWPNGNNT 495
Db 419 AKIDYLDKVKQPSFVKDRAMKREVEEFKIRINALVSKALKVPEEGIMQDGTWPNGNNT 478
QY 496 RDHPGMIQVFLGHSGGLDTEGNELPRLVYVSRKPRGFQHHKKAGAMNALVRVSAVLTNG 555
Db 479 RDHPGMIQVFLGHSGGLDTEGNELPRLVYVSRKPRGFQHHKKAGAMNALVRVSAVLTNG 538
QY 556 QYMLNLDCHYTNNSKAVREACFLMDPNLGNQVQVQFPQDFDGDIDRNDRYANRNTVFF 615
Db 539 PFILNLDCHYTNNSKALREACFLMDPNLGNQVQVQFPQDFDGDIDRNDRYANRNTVFF 598
QY 616 DNLRLGLDGIQGVYVGTGVNRYAIYGYEPPKRA--KKPGFLASLCGGKKKASKSKR 673
Db 599 DNLRLGLDGIQGVYVGTGVNRYAIYGYEPPKRA--KKPGFLASLCGGKKKASKSKR 658
QY 674 SSDKSKNHVDSVPVFNLEIDIEGVGAGFDDKSVLMQMSLEKRFQGSAAVFASTL 733
Db 659 ESDKKKSGRHTDSTVPVFNLDIDIEGVGAGFDDKSVLMQMSLEKRFQGSAAVFASTL 718
QY 734 MEYGVQPSQTESLILKEAHLVISCYEDKSEWGTGIEGWTGYSVTEDILTGFKMHARGWR 793
Db 719 MENGVQPSQTESLILKEAHLVISCYEDKSDWMEIGWYISVTEIDILTGFKMHARGWR 778
QY 794 SYCYMKPFAKGSAPINLSRLNQVLRWALGSEVILFSRHCPIMYGYGRLLKFLERFAY 853
Db 779 SYCYMKPFAKGSAPINLSRLNQVLRWALGSEVILFSRHCPIMYGYGRLLKFLERFAY 838
QY 854 INTTYPLSLPLVYCIIPALCLLTGKFMPEISNLASTWFTALFSLFATGILEMRWS 913
Db 839 VNTTYPIITSIPLLMYCTLPVCLFTNQFIIPQISNIASLWFLSLFSLFATGILEMRWS 898
QY 914 GVGIDSWRNEQFWITGGTSAHLFAVFOGLLKVLGIDNFTVTSKANDEEGDFAELYMF 973
Db 899 GVGIDSWRNEQFWITGGTSAHLFAVFOGLLKVLGIDNFTVTSKANDEEGDFAELYMF 958
QY 974 KWTLLIPPTLLIINMGVWAGTSVAINSYGQSWGPGFKGLFFAFWIVHLYPFLKGLM 1033
Db 959 KWTLLIPPTLLIINMGVWAGTSVAINSYGQSWGPGFKGLFFAFWIVHLYPFLKGLM 1018
QY 1034 GRONRPTIVYVAVLLASTFSLWVRVDPFTRLAGPNQTCGINC 1080
Db 1019 GRONRPTIVYVAVLLASTFSLWVRVDPFTRLAGPNQTCGINC 1065

RESULT 8
ID ABB93356
XX ABB93356 standard; Protein; 1065 AA.
AC ABB93356;
XX
DT 31-MAY-2002 (first entry)
XX

DE Herbicidally active polypeptide
XX Herbicidal; plant; agricultural
KW Herbicidal; plant; agricultural
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO200210210-A2.
PN 07-FEB-2002.
XX 28-AUG-2001; 2001WO-EP09892
PR 28-AUG-2001; 2001WO-EP09892
XX (FARB) BAYER AG.
PA Tietjen K, Weidner M;
PI WPI; 2002-269010/31.
DR Identifying plant target proteins comprising aligning and comparing nucleic acid sequences from plant with nucleic acid sequences from non-plant organisms -
XX Claim 5; SEQ ID NO 2567; 26
PS The invention relates to identifying target proteins comprising aligning and comparing nucleic acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using a suitable search algorithm, the E-value of 3 or greater by a factor of 3 or more, the sequences are selected. The sequences are useful for identifying herbicides.
XX Query Match 78.4%;
SQ Best Local Similarity 78.6%;
Matches 839; Conservative

QY 19 CQICADGLGTTLDGVDFTACDV 78
Db 20 CQICADGLGTTLDGVDFTACDV 79
QY 79 GEGDDTDADDGS-DEFNYPASG 137
Db 80 GKDDEGLADEGIVENYPP--- 127
QY 138 KYDSGEIPRGVYPSVTNSQ-MS 195
Db 128 -----EVSHNHLPRLTSRQDTS 182
QY 196 FSGSIGNVANKRVDGKWKQDK 255
Db 183 DPGVLGNVANKRVDGKWKQDK 238
QY 256 ETRQPLSRKVPATSKINPYRM 315
Db 239 EARQPLSRKVPATSKINPYRM 298
QY 316 ALSWILDQFPKWPINRETYLDR 375
Db 299 ALSWILDQFPKWPINRETYLDR 358
QY 376 VLSILAVDYPVKVSCYVSDGAS 435
Db 359 VLSILAVDYPVKVSCYVSDGAS 418
QY 436 QKIDYLDKVKQPSFVKDRAMKRE 495

SEQ ID NO 2567.
herbicide.
s for herbicidally active compounds, g nucleic acid or amino acid sequences amino acid sequences from non-plant organisms -
Sequence Listing: English.
The invention relates to identifying target proteins comprising aligning and comparing nucleic acid or amino acid sequences from plant sequences from non-plant organisms using a suitable search algorithm, the E-value of 3 or greater by a factor of 3 or more, the sequences are selected. The sequences are useful for identifying herbicides. The identified modulators are

are 4530; DB 23; Length 1065;
d. No. 0;
Matches 104; Indels 26; Gaps 8;
CRPCYEHKKEGTQACLOCKTKYKRRHSGPAIR 78
CRPCYEHKKEGTQACLOCKTKYKRRHSGPAIR 79
SOKIADRMWRMNTGGNGVHPKYSGEIGLS 137
EKISERMLGHLTRGKEEGEPQYDK----- 127
ASPDHMMSPGNGISRRAPPY-VNHSPNPSRE 195
ASPERLSVSTIAGGKRLPYSSDVNQSPNRRIV 182
MTNGTSTAPSEGRATDIDASTEYNMEDALLND 255
V-----STQASE-RGGVDIDASTDILLADEALLND 238
VLSVLSIFLHYRLTNPNVFNAYPLWLLSVICIEWF 315
VLSVLSIFLHYRLTNPNVFNAYPLWLLSVICIEWF 298
HYDEGEPSQLAAVDIFVSTVDPLKEPPIVANT 375
HYDEGEPSQLAAVDIFVSTVDPLKEPPIVANT 358
FDALAEETSEFARKWVPFVKKYDIEPRAPERYFC 435
FDALAEETSEFARKWVPFVKKYDIEPRAPERYFC 418
FKIRINALVSKALKVPEEGIMQDGTWPNGNNT 495

PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	28-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	26-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	28-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	26-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	28-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	26-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.

Query Match 72.9%; Score 4209.5; DB 21; Length 959;

Best Local Similarity 80.4%; Pred. No. 0;

Matches 783; Conservative 80; Mismatches 92; Indels 19; Gaps 6;

Qy	111	MRSRMNTGGSGNVGHPKYDSGEIGLSKYDSGEIPRGVPSVTNSQ-MSGEIPGASPDHH	169
Db	1	MLGWHLTRGKGEEMGEPOYDK-----EVSHHLPRLTSRQDTSGEFSNASPERL	49
Qy	170	MMSPGTGNISRAPFPY-VNHSPNPSREFSGSIGNVANKERVGDGKMKQDKGAIPMTNGTS	228
Db	50	SVSSTIAGGKRLPYSSDVNQSNRRIVDPVGLGNVANKERVGDGKMKQDKNTGFPV---ST	106
Qy	229	IAPSEGRAATDIDASTEYNMEDALLNDETROPLSRKYPFIASSKINPYRMVIVLRLVLSI	288
Db	107	QAASE-RGVVDIDASTDILADEALLNDEARQPLSRKYSIPSSRINPYRMVIMLRVLVLC	165
Qy	289	FLHYRLNPVRNAYPLWLLSVICEIFEALSILDOFPKFPINRETLDRLALRYDREGE	348
Db	166	FLHYRINPVENAFALWLSVICEIFEALSILDOFPKFPVNRETLDRLALRYDREGE	225
Qy	349	PSQLAAVDIFVSTVDPLKEPPITANTVLSILAVDYPVKVSCYSDDGASMLTFDALAE	408
Db	226	PSQLAAVDIFVSTVDPLKEPLTANTVLSILAVDYPVKVSCYSDDGAAMLSFSLAE	285
Qy	409	TSEFARKWVPVKYDIEPRAPEYFCQIDYLDKQVQPSFVKDRRAKREYEFKIRIN	468
Db	286	TSEFARKWVPFCKYISIEPRAPEYFAAKIDYLDKQVTSFVKDRRAKREYEFKIRIN	345
Qy	469	ALVSKALKVPEEGIMODGTPWPGNTRDHPGMIQVFLGHSGGLDTRGNELPLRVVYSRE	528
Db	346	ALVSKALKCPEGWVMDGTPWPGNTRDHPGMIQVFLGQNGGLDAGNELPLRVVYSRE	405
Qy	529	KRPGFQHHKAGAMNALRVSAVLTNGOYMLNLCDDHYINNSKAVREAMCFMDPNLGPQ	588
Db	406	KRPGFQHHKAGAMNALRVSAVLTNGPFIINLDCDDHYINNSKALREAMCFMDPNLKGQ	465
Qy	589	VCYVQFPQRFQDIDRNDRYANRNTVFFDINLRGLDGTGGVYVYCTGCVFNRTALYGEPP	648
Db	466	VCYVQFPQRFQDIDRNDRYANRNTVFFDINLRGLDGTGGVYVYCTGCVFNRTALYGEPP	525
Qy	649	IKA--KKPGFLASICGGKKKSKRSOKKSNKHWDSSVPVFNLEDIEEGVEGAGFD	706
Db	526	IKVHKKPSLLSKLGGSRKKNKAKKESDKKSGRHTDSTVPVFNLEDIEEGVEGAGFD	585
Qy	707	DEKSVLMSQMSLEKRRFGQSAAFVASTLMYEGVPGQSTPESLLKEA IHVTS CGYEDKSEW	766
Db	586	DEKALLMSQMSLEKRRFGQSAFVASTLMENGVPPSPATPENLLKEA IHVTS CGYEDKSDW	645
Qy	767	GTEIGWIYGSVTE DILTGFKMHARGWSVVCMPKRPAPKGSAPINLDRNLQVLRWALGS	826
Db	646	GMEIGWIYGSVTE DILTGFKMHARGWSIYCMPLKPAKGSAPINLDRNLQVLRWALGS	705
Qy	827	VEILFSRHCPWYGGGRKFLERFAYINTTIYPLTSLPLIVYICILPAICLLTGKFTMPE	886
Db	706	VEILFSRHCPWYGNGLRFLERFAYVNTTIYPTISPLMYCTLPVCLFTNQFIIPQ	765
Qy	887	ISNLASITWFIALFTSIFATGILEMRWSGVGIDEWRNEQFWVIGGISAHFAVFPQGLLK	946

Db 766 ISNIASIFLSLFSIFATGILEMRWSGVGIDEMWRNEQFWIGVSAHLFAVFOGILKV 825

QY 947 LAGIDNFTVTSKANDEEGDFAELYFKWTTLLIPTTTLLINMGVWAGTSVAINSGYQ 1006

Db 826 LAGIDNFTVTSKANDEEGDFAELYFKWTTLLIPTTTLLINMGVWAGTSVAINSGYQ 885

QY 1007 SWGPLEGKLFFAFWVTHLYPFLKGLMGQRNRPRTTIVIVAVLLASIFSLWVRVDPFTT 1066

Db 886 SWGPLEGKLFFAFWVTHLYPFLKGLMGQRNRPRTTIVIVAVLLASIFSLWVRVDPFTS 945

QY 1067 RLAGPNIOQCGINC 1080

Db 946 RVTGPDILECGINC 959

RESULT 11

AY84110

ID AAY84110 standard; Protein; 1074 AA.

AC AAY84110;

XX

DT 03-JUL-2000 (first entry)

XX

DE Amino acid sequence of a maize cellulose synthase.

XX

KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

KW transgenic plant; plant breeding marker.

XX

OS Zea mays.

XX

PN WO200009706-A2.

XX

PD 24-FEB-2000.

XX

PF 16-AUG-1999; 99WO-US18760.

XX

PR 17-AUG-1998; 98US-0096822.

XX

PA (PION-) PIONEER HI-BRED INT INC.

XX

PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX

DR WPI; 2000-224343/19.

DR N-PSDB; AA299500.

XX

PT New genes which encode maize cellulose synthase polypeptides in plants

PT useful for modulating the expression of cellulose synthase in plants

PT and to produce transgenic plants expressing the novel protein

XX

PS Claim 15; Page 110-113; 119pp; English.

XX

CC The present sequence represents a maize cellulose synthase polypeptide.

CC The cellulose synthase can be used for the improvement of stalk quality

CC for improved stand or silage. It also provides an increased concentration

CC of cellulose in the pericarp, hardening the kernel and improving its

CC handling ability. The sequences are used to produce transgenic plants

CC and seeds expressing the cellulose synthase. The polynucleotide is

CC used for modulating, preferably increasing, the level of the synthase

CC in a plant cell. The plants are preferably monocots. The polynucleotide

CC is also used as a probe or primer in the detection quantitation or

CC isolation of gene transcripts. The probes are useful in detecting

CC deficiencies in the level of mRNA in screenings for desired transgenic

CC plant, for detecting mutations in the gene, for monitoring upregulation

CC of expression or changes in enzyme activity in screening assays of

CC compounds, for detection of any number of allelic variants of the gene,

CC or for use as molecular markers in plant breeding programs. The

CC isolated nucleic acids of the present invention can also be used for

CC recombinant expression of their encoded polypeptides or for use as

CC immunogens in the preparation and/or screening of antibodies. The

CC proteins can be employed in assays for enzyme agonists or antagonists

CC of enzyme function or for use of immunogens or antigens to obtain

CC antibodies specifically immunoreactive with a protein.

SQ	Sequence	1074 AA:
QY	Query Match	69.1%;
Db	Best Local Similarity	68.4%;
QY	Mismatches 749; Conservative	1
Db		
QY	2 DGADA-LKSGRGAGDVCOIC	69.1%;
Db	23 DGADVPAPKTKSANGQVCOIC	68.4%;
QY	61 CLOCKTKYKRRHRSIPAIRGEG	69.1%;
Db	83 CPOCKTRYKRGKSPRVHGGD-	68.4%;
QY	120 GSGNVGHPKY---DSGEIGLS	69.1%;
Db	120 --GNGKGPWQLQGDADLSS-	68.4%;
QY	173 PTGN-ISRRAFPFYVNHSPNS	69.1%;
Db	174 PTSSYVDPSPVPV--RIVDPS	68.4%;
QY	231 PSEGRAATDIDASTEYNMEDAL	69.1%;
Db	227 -KYPEARGDME-GTSGNGEDMQ	68.4%;
QY	291 HYRLTNPVRNAYPLWLLSVICB	69.1%;
Db	285 QYRISHPVRNAYGLWLSVICB	68.4%;
QY	351 QLAADVIFVSTVDPLKEPPVIT	69.1%;
Db	345 QLAPIDVFVSTVDPLKEPPLIT	68.4%;
QY	411 EFARKWVPFVKYKDYIEPRAPEE	69.1%;
Db	405 EFARKWVPFCKKHNIIEPRAPEE	68.4%;
QY	471 VSKALKVPEEGTWQDGTWPG	69.1%;
Db	465 VAKAQVPEEGWTMADGTWPG	68.4%;
QY	531 PGFQHHKAGAMNALVRVSAVI	69.1%;
Db	525 PGFQHHKAGAMNALVRVSAVI	68.4%;
QY	591 YVQFPORFDGIDRNDYRNRNT	69.1%;
Db	585 YVQFPORFDGIDLHRYRNRNT	68.4%;
QY	651 AK--KFCFLASLCGGKKASKS	69.1%;
Db	645 EADLEPNIVVKSCGRRK-RKN	68.4%;
QY	709 KSVLMQSMLEKRFSGSAFVA	69.1%;
Db	702 RSVLMQSMLEKRFSGSPIFA	68.4%;
QY	769 EIGWIGSVTEDIITGFKMHAR	69.1%;
Db	762 EIGWIGSVTEDIITGFKMHAR	68.4%;
QY	829 ILFSRHCPWYGGRLKFLER	69.1%;
Db	822 ILLSRHCPIWYNGRLKLER	68.4%;
QY	889 NLASIFIALFIFATGILEM	69.1%;
Db	882 NYAGMFFILLFASFATGILEI	68.4%;
QY	949 GIDTNTVTYSKANDEGDFAEI	69.1%;
Db	942 GIDTNTVTYSKANDEGDFAEI	68.4%;

are 3993.5; DB 21; Length 1074;
ad. No. 0;
Mismatches 153; Indels 59; Gaps 20;
STTLDDGVFTACDVCVCFPPVCRPCYEHKKEGTQA 60
VSATGDVFAVACNECAFVCRPCYEHKKEGNOC 82
DD-GSDFNTPASCTEDQOKIADMRMSWRNWTG 119
DLDNEFN-----KQ----- 119
GGEIPRGVFSVTN-SOMSGEIPGASPDH-MMS 172
---PHRIRPLTSGQISGEIPDASPDHRSIRS 173
SGSIGNVAWKERVGDGKWKQDKGAIMPTNGTSIA 230
SYGLNSVDWKERVESWRVKQDKNMLQVTN----- 226
RQPLSRKVPTASSKINPYRMVIVLVLVLSIFL 290
RLPLSRIVPFISSNOLNLYRIVITLRLILGFF 284
SWILDQFPKWFPIINRETYLDRLALRDREGEPS 350
SWILDQFPKWFPIINRETYLDRLALRDREGEPS 344
SILAVDYPVKVSCYVSDDGASMLTDLAETS 410
SILAVDYPVKVSCYVSDDGASMLTDLAETS 404
SIDYLDKQVOPSFVKDRANKREVEEKIRINAL 470
SIDYLDKQVOPSFVKDRANKREVEEKIRINAL 464
HPGMIQVFLHSGGLDTEGNEPLRLVYVVSREKR 530
HPGMIQVFLHSGGLDTEGNEPLRLVYVVSREKR 524
MLNLDCHYINNSKANVREACFLMDPNLGPQVC 590
LLNVDCDHVFNSSKALREACFMMDPALGRKTC 584
NLRLGDIQGPVYVGTGCVFNRTAITYGPEPIK 650
NMKGLDGIQGPVYVGTGCCFNRLQALGYDPLT 644
SDKKKSNKHVDSSVVPFNLIEDIEGVEGAGFDE 708
SQSRIMKRTRESSAPIFNMEDIEGIE--GYEDE 701
SYGGVPOSSPESLLKKAHIVISGVEDKSWGT 768
YGGIPPTNPASLLKKAHIVISGVEDKSWGT 761
YCMKRPAPKGSAPINLSDRNLQVLRWALGSVE 828
YCMPPRCFKGSAPINLSDRNLQVLRWALGSVE 821
YTTIYPLTSLPLVYICLPAICLTLTKGFIMPEIS 888
YTTIYPLTSLPLVYICLPAICLTLTKGFIMPEIS 881
YDIEWNRNEQFWIGTSIAHLFAVFOGLLVLA 948
YDIEWNRNEQFWIGTSIAHLFAVFOGLLVLA 941
YTTLLIPPTTLLIINMGVWAGTSVAINSGYQSW 1008
YTTLLIPPTTLLIINMGVWAGTSVAINSGYQSW 1001

[REDACTED]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2003, 16:37:20 ; Search time 50 Seconds
(without alignments)
2076.504 Million cell updates/sec

Title: US-09-900-237-30
Perfect score: 5778
Sequence: 1 MDGADALKSRHGAGDVCG.....VDFPTRLAGPNIQTCGINC 1080

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4507	78.0	1065	2 T52054	cellulose synthase
2	3963.5	68.6	1081	2 T05351	cellulose synthase
3	3801	65.8	1026	2 T51579	cellulose synthase
4	3778.5	65.4	1065	2 F84649	probable cellulose
5	3666.5	63.5	1084	2 T08583	cellulose synthase
6	3593	62.0	1081	2 T52028	cellulose synthase
7	3579	61.9	1088	2 H84604	probable cellulose
8	3477	60.2	974	2 T10797	cellulose synthase
9	3302.5	57.2	958	2 T04870	cellulose synthase
10	2816	48.7	685	2 T10800	cellulose synthase
11	2315	40.1	1181	2 D86157	hypothetical prote
12	2257	39.1	1111	2 T05646	hypothetical prote
13	2220.5	38.4	1145	2 T51546	cellulose synthase
14	2212.5	38.3	1036	2 D84741	probable cellulose
15	2163.5	37.4	979	2 C86446	probable cellulose
16	1941.5	33.6	583	2 T02209	cellulose synthase
17	1094	18.9	757	2 T02561	probable cellulose
18	1076	18.6	712	2 T02552	cellulose synthase
19	1057.5	18.3	755	2 T02553	cellulose synthase
20	1034.5	17.9	748	2 T02560	cellulose synthase
21	983	17.0	689	2 T08918	hypothetical prote
22	963.5	16.7	686	2 T08919	hypothetical prote
23	959.5	16.6	727	2 T08920	hypothetical prote
24	943.5	16.3	828	2 E71417	hypothetical prote
25	872.5	15.1	710	2 B71417	hypothetical prote
26	430	7.4	693	2 AF2275	cellulose synthase
27	289.5	5.0	759	2 D70422	cellulose synthase
28	271.5	4.7	322	2 T12093	TGACG-motif bindin
29	270	4.7	326	2 T08591	TGACG-motif bindin

ALIGNMENTS

RESULT 1

T52054
cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C:Accession: T52054
R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C
Science 279, 717-720, 1998
A:Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A:Reference number: Z13745; MUID:98111412; PMID:9445479
A:Accession: T52054
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1065 <ARI>
A:Cross-references: EMBL:AF027174; PIDN:AAC39336.1
C:Genetics:
C:Note: Ath-B
C:Function:
A:Description: EC 2.4.1.-; cellulose synthase [validated, MUID:98111412]; Involved in
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match	78.0%	Score	4507;	DB	2;	Length	1065;
Best Local Similarity	78.4%;	Pred. No.	0;				
Matches	836;	Conservative	98;	Mismatches	107;	Indels	26;
Gaps	8;						
QY	19	COICADGLTLDGDFVTACDVCFRPCYERHERKEGTQACLOCKTKYKRRGSPAIR	78				
DB	20	COICSDNVGKTVGDRFVACDICSFPVPCYERKDGNSQCPQCKTRYKRLKGSIPAIP	79				
QY	79	GEGDDTDADGGS-DFNYPASGTEDQKKTADRMRSRMNTGGSGNVGHPKYDSGEIGLS	137				
DB	80	GDKDEGLADEGTVEFNYP-----QREKISERMLGHLTRGKEGMEGPQYDK-----	127				
QY	138	KYDSEIPRGVPSVTSNQ-MSGEIPGASPDHMHMSPTGNISRRAPFPY-VNHSNPSPRE	195				
DB	128	-----EVSHNHLPLRTSQDTSQDTSGEFSAASPERLSVSSTIAGGKRLPYSDVQNSPNRIV	182				
QY	196	PSGSGTGNVANKERVDGNKMKQDKCAIPMTNGTSTIAPSEGRAATDIDASEYNMEDALLND	255				
DB	183	DVPGUGNVANKERVDGNKMKQKERTNPV---STQAAE-RGGVDIDASTDILADEALLND	238				
QY	256	ETROPLSRKVPITASSKINPYRMVILRLVLVLSIFLHYRLNPVRNAYPLWLLSVICEIWF	315				
DB	239	EARQPLSRKVSIPSSRNIPYRMVIMLRVLVILCLHYRITNPVPNFAFWLSVICEIWF	298				
QY	316	ALSWILDOPKFWPPIRNRTYLDRLALRYDRGEPSQAAVDIFVSTVDLPKEPIVTANT	375				
DB	299	ALSWILDOPKFWPPIRNRTYLDRLALRYDRGEPSQAAVDIFVSTVDLPKEPIVTANT	358				
QY	376	VLSILAVDYPVDKVCSCYCYSDDGASMLTFDALAETSEFARKWVPFKKYDIEPRAPEYFC	435				
DB	359	VLSILAVDYPVDKVCSCYCYVDDGAAMLSFESLAETSEFARKWVPFKKYDIEPRAPEYFA	418				

probable polysacch
cellulose synthase
probable cellulose
hypothetical prote
hypothetical 101.6
cellulose synthase
cellulose synthase
cellulose synthase
TGACG-motif-bindin
probable cellulose
bcsA protein - Ace
probable cellulose
cellulose synthase
glycosyltransferas
hypothetical prote
beta 1,3 glucan sy

Qy	436	QKIDYLDKQVPSFVKDRAMKREYEEFKIRINALYSKALKYPERGWINQDTPPGNNT	495
Db	419	AKIDYLDKQVQTSFVKDRAMKREYEEFKIRINALYSKALKYPERGWINQDTPPGNNT	478
Qy	496	RDPGMIQVFLHSGGLDTFEGNELPRLVVSREKRGPGFQHHKKAGAMNALVRVSALVING	555
Db	479	GDHPGMIQVFLGQGLDAEGNELPRLVVSREKRGPGFQHHKKAGAMNALVRVSALVING	538
Qy	556	OYMLNLDCHYINNSKAVREACMFMDPNLGPQVYVQPPQDFGIDRNDRYANRNTVFF	615
Db	539	PFILNLDCHYINNSKALREACMFMDPNLGVQVYVQPPQDFGIDRNDRYANRNTVFF	598
Qy	616	DINLRGLDGIQGVVYVGTGCVFNRTAIYGEYPIKA--KKPGFLASLCGGKKKASKKR	673
Db	599	DINLRGLDGIQGVVYVGTGCVFNRTAIYGEYPIKVKHKKPSLLSKLGGSRKKNSKAKK	658
Qy	674	SSDKKSNKHVDSSVPVFNLEDIEEGVEGAGDEDEKSVLMSQMSLEKRFQSGAAVASTL	733
Db	659	ESDKKSGRHTDSTVPVFNLDIEEGVEGAGDEDEKALMSQMSLEKRFQSGAVASTL	718
Qy	734	MEYGVPPQSGTSPSLKKEAIHVISCYEDKSEWGTGIGWYGSVTEIDITGFKMHARGWR	793
Db	719	MENGVPSPATPENLKEAIHVISCYEDKSDWGEIGWYGSVTEIDITGFKMHARGWR	778
Qy	794	SVYCMKPRFAFKSAPINLSRDLNQVLRWALGSVELFSRHCPWYGGRLKFLERFAY	853
Db	779	SIYCMPLKFAFKSAPINLSRDLNQVLRWALGSVELFSRHCPWYGGRLKFLERFAY	838
Qy	854	INTIYPLTSLPVLVYCIPLAICLTGKFIPEISNLASITWFTALFLSFATGILEMRWS	913
Db	839	VNTIYPTISPLMYCTLLAVCLFTNQFIIPQISNIAISWFLSLFSLFATGILEMRWS	898
Qy	914	GVGIDEMWNEQFWTGGISAHLFVAFQGLLKLVLGADITNFTVTSKANDEGDFAEIYMF	973
Db	899	GVGIDEMWNEQFWTGGISAHLFVAFQGLLKLVLGADITNFTVTSKANDEGDFAEIYLF	958
Qy	974	KWTTLLIPTPTILLINMGVAGTSYAINSGYQSGMPLFGKLFFAFWVIVHLVLPKLGLM	1033
Db	959	KWTTLLIPTPTILLINVLGVAGTSYAINSGYQSGMPLFGKLFFAFWVIVHLVLPKLGLM	1018
Qy	1034	GRONRPTIYVWVLLASIFSLWVRVDPFTTLRAGNIOTCGINC	1080
Db	1019	GRONRPTIYVWVLLASIFSLWVRVDPFTTLRAGNIOTCGINC	1065
RESULT 2			
T05351			
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana			
N:Alternate names: protein F8B4.110			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999			
C:Accession: T05351			
R:Revan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De			
ewes, H.W.; Mayer, K.F.X.; Schueller, C.			
submitted to the Protein Sequence Database, February 1999			
A:Reference number: Z15409			
A:Accession: T05351			
A:Molecule type: DNA			
A:Residues: 1-1081 <BEV>			
A:Cross-references: EMBL:AL034567			
A:Experimental source: cultivar Columbia; BAC clone F8B4			
C:Genetics:			
A:Gene: RSW1			
A:Map position: 4			
A:Intons: 27/3; 93/1; 150/3; 169/1; 252/2; 341/2; 456/3; 502/3; 544/3; 704/1; 76			
A:Note: F8B4.110			
C:Keywords: glycosyltransferase; hexosyltransferase; P-loop			
Query Match 68.6%; Score 3963.5; DB 2; Length 1081;			
Best Local Similarity 68.0%; Pred. No. 3.le-282;			
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;			

Qy	2	DGDADALSGRHGAGDYCOICA	61
Db	25	DGGTKPLKNNM--GQICQICG	81
Qy	62	LQCKTKYKRRHSGSPAIRGEED	121
Db	82	POCKTRFRHRHSGSPRVEGDEE	133
Qy	122	GNVGHKPYDSGEIGLSKYD--	178
Db	134	---SSRHESOPILLTHGHTW	184
Qy	179	RRAPFPVYNNHSPNREF--SS	237
Db	185	PRQPVV--RIVDPSKDLNSYG	238
Qy	238	TDIDASTEYNMEDALLNDETRO	297
Db	239	-EIE-GTSGNGEELQMADDTRI	296
Qy	298	VRNAYPLWLLSVICELWFAFSK	357
Db	297	VKNAYPLWLLSVICELWFAFSK	356
Qy	358	FVSTVDPDLKEPPIYVANTVLS	417
Db	357	FVSTVDPDLKEPPIYVANTVLS	416
Qy	418	PFVKYDIETPRAPFIFYCOKIE	477
Db	417	PFCKKFNIEPRAPFIFYCOKIE	476
Qy	478	PEEGWIMODGTPWPGNNTROH	537
Db	477	PEEGWIMODGTPWPGNNTROH	536
Qy	538	KAGAMNALVRVSALVTNGQYM	597
Db	537	KAGAMNALVRVSALVTNGAYL	596
Qy	598	FDGIDRNDRYANRNTVFFDIN	655
Db	597	FDGIDRNDRYANRNTVFFDIN	656
Qy	656	FLA-SLCCGKKKASKKRSS	714
Db	657	IIVKSCCGSRKKGSKKKNY	714
Qy	715	QMSLEKRFQSGAAFAVASTLME	774
Db	715	QRSVEKRFQSGSPVFIATFME	774
Qy	775	GSVTEDILTGFKMHARGWRSW	834
Db	775	GSVTEDILTGFKMHARGWIST	834
Qy	835	CPLWYGYGGRKLFLERFAYIN	894
Db	835	CPIWYGYHGRLLRLERAYIN	894
Qy	895	FIALFLSIFATGILEMRWSGV	954
Db	895	FILLFISIAVTGILELRWSGV	954
Qy	955	TVTSKANDEGDFAEIYMFKA	1014
Db	955	TVTSKATDEGDFAEIYFKK	1014
Qy	1015	LFFAFWVIVHLVLPKLGLMGS	1072
Db	1015	LFFALWVIAHLVLPKLGLLGE	1070

Qy 348 EPSQLAAVDIEFVTDVPLKEPIPIVNTANTVLSILAVDYPVDKVCYVSDGASMLTFDALA 407
Db 334 EPSQLAPVDVFSVTDVPMKEPPLVTANTVLSILAVDYPVDKVCYVSDGASMLTFEALS 393
Qy 408 ETSEFAKWPVFKVYDIEPRAPPEYFCQKIDYLDKVKQVPSFVKDRAMKREYEEFKIRI 467
Db 394 ETAFESKWWPFCKENIEPRAPPEYFSQKIDYLDKVKQVPSFVKDRAMKREYEEFKIRI 453
Qy 468 NALYSKALKVPEEGWIMQDGTWPNNTRDHPGMIOVFLGHSGLDTEGNEPLRLVYVSR 527
Db 454 NILVAKAQKIEDGWTMEDGTWFGNPRDHPGMIOVFLGHSGLDTEGNEPLRLVYVSR 513
Qy 528 EKRPFGFHHKAGAMNALVRVAVLTNGQYMLNDCDHYINNKAIVREAMCFMDPNLGP 587
Db 514 EKRPFGFHHKAGAMNALIRVAVLTNGAYLLNVDCHYFNNSKAIRKAEAMCFMDPDAIGK 573
Qy 588 QVCYVQFQRFQDIDRNDRYANRNTVFDINLRGLDGIQGVYVGTGCVFNRTALYGEPE 647
Db 574 KCCYVQFQRFQDIDLHRYANRNTVFDINLRGLDGIQGVYVGTGCVFNRTALYGEPE 633
Qy 648 PIKAK--KPGFLASLCGGKKKASKSK--RSSDKKKSNKHVDSSVPVFNLEDIEBEGVAG 704
Db 634 VLTEEDLPNIIVKSCFGRKKGSKKIPNYEDNRSKRSDSNVPLFNMEDIDEDVE--G 691
Qy 705 FDDEKSVLMSQMSLEKRFQSGAAVASTLMEYGGVPOSSTPESLLKKAHIVISCGYEDKS 764
Db 692 YEDEMSLVSKRLEKRFQSGQSVFIATFMEQGLPSTTNPLTLKKAHIVISCGYEAKT 751
Qy 765 EWGTEIGWYGSVTEIDLTGFKMHARGWRSVYCMKPAFKGAPINLSDRNLQVLRWAL 824
Db 752 DWGKEIGWYGSVTEIDLTGFKMHARGWISVYCVSPAFKGSAPINLSDRNLQVLRWAL 811
Qy 825 GSVTEILSRHCPLMWGYGRLEKFERPAYINTIYPLTSLPLLYCYILPAICLLTGKFM 884
Db 812 GSEIELLSRHCPWGYNGRLKLERIAYINTIYPTISIPILAYCMPLPAFCLTNTFII 871
Qy 885 PEISNLASIWTFILFSLFATGILEMRWSGVGIDEMWNEQFVWIGGISAHLFVFOGLL 944
Db 872 PEISNLASLCFMLFASIAAILEKMSDVALEDWNRNEQFVWIGGISAHLFVFOGLL 931
Qy 945 KVLADIDNFVTSKANDEEDGFALYMKWTTLLIPPTTILINMVGVAAGTSVAINSG 1004
Db 932 KVFAGIDNFVTSKASDEEDGFALYVFKWTSLLIPPTTILINMVGIVAGVAINSG 991
Qy 1005 YQSWGPLEGKLFFAFWVHLVPLFLKGLMGRQNRTPTIIVMAVLLASIFSLLMVRVDPF 1064
Db 992 YQSWGPMGLKLLFAFWVVAHLYPFLKGLLGRQNRTPTIIVVWSALLASIFSLLMVRINPF 1051
Qy 1065 TTRLAGPNIQTCGI 1078
Db 1052 -----VSTGTG 1057

RESULT 5
T08583
cellulose synthase (EC 2.4.1.-) catalytic chain - Arabidopsis thaliana
N:Alternate names: protein f22f8.250
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
C:Accession: T08583; T09014
R:Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, May 1999
A:Reference number: 216442
A:Accession: T08583
A:Molecule type: DNA
A:Residues: 1-1084 <BEV>
A:Cross-references: EMBL:AL050351; GSPDB:GN00062; ATSP:T22F8.250
A:Experimental source: cultivar Columbia; BAC clone T22F8
R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H
Science 279, 717-720, 1998
A:Title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A:Reference number: Z13745; M0UD:98111412; PMID:9445479

A:Accession: T09014
A>Status: translated from GB/EMBL
A:Molecule type: mRNA
A:Residues: 1-1084 <ARI>
A:Cross-references: EMBL:AF02717
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:T22F8.250; Ath-A
A:Map position: 4
A:Introns: 27/3; 156/3; 191/1; 2
C:Keywords: cell wall synthesis;
Query Match 63.5%;
Best Local Similarity 62.4%;
Matches 687; Conservative 1
Qy 1 MDGDADA-LKSGRHGAGDVCOI 1
Db 20 INADESARISVQELSQTQOI 1
Qy 60 ACLOCKTKYKRRGSPAIRGEE 1
Db 80 ACPOCKTRYKRIKGSPPV--D 1
Qy 120 GSGNVGHPKYDSGIGLSKYDS 1
Db 136 RGLDLSAP----- 1
Qy 178 SRRAPFYVNHSPNPS----- 1
Db 177 NRVYPAPFTDSSAPPQARSVP 1
Qy 227 TSITAPSEGRAATDIDASTEYNN 1
Db 235 ---GNGRGSNDDD---ELDE 1
Qy 287 SIFLHYRLTPVRNAYPLWLLS 1
Db 288 GLFPHYRLHPVNDAYGLWLTS 1
Qy 347 GEPQLAAVDIFVSTVDPLKE 1
Db 348 GKPSGLAPDVVFVSTVDPLKE 1
Qy 407 AETSEFAKWPVFKKVDIEPE 1
Db 408 SDTAEFAKWPVFCCKFNIEP 1
Qy 467 INALVSKALKVPEEGWTMDQS 1
Db 468 INALVATAQKVPESGWTMDQS 1
Qy 527 REKRPFGFHHKAGAMNALVR 1
Db 528 REKRPFGFHHKAGAMNSLIR 1
Qy 587 POVCYVQFQRFQDIDRNDRY 1
Db 588 KVCYVQFQRFQDIDRNDRY 1
Qy 647 PPIKAKPGFVLAS-----L 1
Db 648 APKKKKPGKTCNCWPKWCC 1
Qy 700 VEGAGFDDERSVLMQSLEK 1
Db 704 VIVPVSNNVKKRSEATQLKEK 1
Qy 760 YEDKSEWTEIGWYGSVTEID 1
Db 764 YEDKTEWKEIGWYGSVTEID 1
Qy 820 LRWALGSVEILFSRHCPMWY 1

J
ID:g2827140; PIDN:AAC39335.1; PID:g2827141
bia
343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3
osyltransferase; hexosyltransferase
are 3666.5; DB 2; Length 1084;
ad. No. 1.9e-260; Indels 59; Gaps 16;
Mismatches 191;
GTTLDGVDFTACDVCRRFPVPCRCYEHKREGTQ 59
ELTVSSSELFVACNECAFFVPCRCYEVERRGNQ 79
SADGSDFNYPASGTEDOKQKIADMRSMRNTG 119
SEEDIDDDLEYEFDHGM--PERAAALSSRLNTG 135
RGYPVSVTNSQMSGEIPGASPDHMMSP--TGNI 177
SSQIPLLTYCEDADAM--YSORHALIVPSTGYG 176
REFSGSIGNVAKKERVDCGKMKQ--DKGAIPWTNG 226
AEY--GYGSAVAKMORMEVKKRQSEKQIVIKHEG 234
LNDETQPLSRKVPPIASSKINPYRMVILRLVVL 286
WMDEGROPLSRKLPTRSSRINPYRMILCLRIL 287
LWFALSWILDOPFKWFPINRNTYLDRLALRYDRE 346
LWFAVSWILDOPFKWYPIERETYLDRLSRLEKE 347
LANTVLSILAVDYPVDKVCYVSDGASMLTFDAL 406
LANTVLSILAVDYPVDKVCYVSDGASMLTFEAL 407
LFCQKIDYLDKVKQVPSFVKDRAMKREYEEFKIR 466
LFSQKMDYLDKVKVHPAFVRRRRAMKRDYEEFKVK 467
LNNTRDHPGMIOVFLGHSGLDTEGNEPLRLVYVS 526
LNNVRDHPGMIOVFLGHSGRVTDGNEPLRLVYVS 527
LNGQYMLNDCDHYINNKAIVREAMCFMDPNLGP 586
LSNAPYLLNVDCDHYINNKAIVRESCFMDPQSG 587
LYFFDINLRGLDGIQGVYVGTGCVFNRTALYGYE 646
LYFFDINLRGLDGIQGVYVGTGCVFNRTALYGYE 647
LYFFDINLRGLDGIQGVYVGTGCVFNRTALYGYE 647
LYKASKSKRSDDKKSNKHVDSSVPVFNLEDIEEG 699
LYK--SKTKAKDKKTKNTK--ETSKQIHALENVDEG 703
LYAAFAVASTLMEYGGVPOSSTPESLLKKAHIVISCG 759
LYSPVFAVAVLQNGVPRNAPSACLLREALQVLSG 763
LYKMHARGWRSVYCMKPAFKGAPINLSDRNLQV 819
LYKMHCHGWRVYCMKPAFKGAPINLSDRNLQV 823
LYKFLERFAINTIYPTISIPILAYCMPLPAICLLT 879

```
|||||
Db 824 LRWALGSVEIFLSRHCPITWYGGGLKWLRFYSINSVVPWTSPLPIVYCSLPVACLLT 883
QY 880 GKETMPEISNLASLWFTALFSLFATGILEMRSGVGYDDEWNEOFWVGTSIAHLFAV 939
Db 884 GRFIVPEISNLAGLFLMDFISIAVTAIGLEMGVGGVDDWNRNEOFWVGASSHLFAL 943
QY 940 FQGLKVLGADTNTFTVTSKANDREGDAELMYFKWTTLLIPPTTLLIINMWGVVAGTSY 999
Db 944 FQGLKVLGAVNTFTVTSKAA--DGAFFSELYIFKWTLLIPPTTLLIINIGVGVSD 1002
QY 1000 AINSYQSGWGLPGLKLEFAFWIVHLYPFLKGLMGQRNPTTIVVAVLLASIFSLWV 1059
Db 1003 AINSYQSGWGLPGLKLEFAFWIVHLYPFLKGLMGQRNPTTIVVAVLLASIFSLWV 1062
QY 1060 RYDPTTLAGPNQTQGCINC 1080
Db 1063 RVNPFVAK--GGPVLEICGLNC 1082

RESULT 6
t52028
cellulose synthase [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: t52028
R.Joshi, C.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z25890
A:Accession: t52028
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1081 <JOS>
A:Cross-references: EMBL:AF062485; PIDN:AAC29067.1

Query Match 62.0%; Score 3583; DB 2; Length 1081;
Best Local Similarity 61.1%; Pred. No. 2.6e-254;
Matches 676; Conservative 162; Mismatches 198; Indels 72; Gaps 15;

QY 1 MDGADADA-LKSGRGAGDVCOICADGLTLDGDVFTACDVCRFPVCRPCYEHREGTQ 59
Db 17 INADENARISVQELSGTQCIDEIELTVDGEPFVACNECAFPVCRPCYEHREGNQ 76

QY 60 ACLOCKTKYKRRHSGPAIRGEGDDTDADGSDFNYPASGTE-DOKOKIADRMRSWRMNT 118
Db 77 ACPOCKTRFKRLKSGPRVGEDEDDIDLDNEPEYCNNGIGFDQ--VSEGMSISRRNS 133

QY 119 GSGNVGHPKYDGSGEIGLSKYDSCGEIPRGYVPSVTNSQMSGEIPGASPDHH----- 169
Db 134 -----GFPQSD-----LDSAP-PGSQLPLTYGDEVEI---SSDRHALIVPPSLG 175

QY 170 -----MMSPTGNISRRAPFPYYNHSFNPSPREFSGSIGNVANKERVVDGKMKQDKG 219
Db 176 GHGNRVHPVLSLDPVTAHRRLMVPQDKLAVY-----GVGSVANKDRMEWKRKQNE- 227

QY 220 AIPWTNGTSIAPSGRAATDIDASTEYNNMEDALLNDETROPLSKRPVIASSKNIPRMVI 279
Db 228 -----KIQVVRHGGDPDFDGDADPFM-----MDGROPMSMKIPKISKNIPRMVI 276

QY 280 VLRLVVLGIFLHYRLTNVRNAYPLWLLSVICEIWFALSILWLDQFPKWFFPINRETYDLRL 339
Db 277 VLRLVILGFFHYRLILHPVKDAYALWLLSVICEIWFALVSVWLDQFPKWYPIERETYLRL 336

QY 340 ALRYDREGPEPQLAAVDIFVSTVDPLKEPPIVTVANTVLSILAVDYPVDKVCSTVSDGAS 399
Db 337 SLRYEKEGKPSGLSPVDVFSVTDVPLKEPPLITANTVLSILAVDYPVDKVCSTVSDGAA 396

QY 400 MLTFDALSETSEFAKWPVKYKIDYKDKQVPSVVKDORRAMKRE 459
Db 397 MLTFDALSETAEFAKWPVKYKIDYKDKQVPSVVKDORRAMKRD 456

QY 460 YEEFKIRINALVSKALVPESGWTMQDGTWPGNSVRDHPGMIQVFLGSDGVRDVENNEL 519
|||||
```

```
Db 457 YEEFKVINALVATAQKVPEDGWTMQDGTWPGNSVRDHPGMIQVFLGSDGVRDVENNEL 516
QY 520 PRLVYVREKRPFGQHHKAGAMNALVRVSAVLNTNGOYMLNLDCHYINNSKAVREAMCF 579
Db 517 PRLVYVREKRPFGDHHKAGAMNSLRVSGVLSNAPYLLNVDCHYINNSKALREAMCF 576
QY 580 LMDPNLGPQVCYVQFQRFQDIDRNDYANRNTVFFEDINIRGLDIOGQVPYVGTGCVFNR 639
Db 577 MNDPQSGKKICYVQFQRFQDIDRNDYANRNTVFFEDINIRGLDIOGQVPYVGTGCVFNR 636
QY 640 TAIYGYEPPKAKKPGFLAS-----LCGGKKKASKSKRSDKKSKNKHVDSSVPVEN 692
Db 637 QALYGFDAKPKKKGPRKTCNCWPKWLLCFGSRNKRKAKTVAADKKKKNR--PASKQIIHA 694
QY 693 LEDIEEGVEGAGFDDKSVLMSQMSLEKRFQSQAAFAVASTLMEYGGVPOSSTPESLLKEA 752
Db 695 LENIEEGRHKVLNVEQSTEAMQMLQKYQSPVFVASARLENGMGGMARNASPAELLKEA 754
QY 753 IHVISCYVEDKSEWGTETGWTGYSVTEDILTGFKMHARGHRSVYVCMKPAFKGSAFINL 812
Db 755 IOVISRGYEDKTEWKGKEIGWYGSVTEDILTGSKMHSHGWRHYCTPKLAAFKGSAPINL 814
QY 813 SDRLNQVLRWALGVSVEILFGRHCPWYGGRLKFLERFAYINTTIYPLTSLPLLYCYIL 872
Db 815 SDRLHQVLRWALGVSVEILFGRHCPWYGGRLKFLERFAYINTTIYPLTSLPLLYCYIL 874
QY 873 PAICLLTGKFTMPEISNLASLWFTALFSLFATGILEMRSGVGYDDEWNEOFWVGIGI 932
Db 875 PAICLLTGKFTMPEISNLASLWFTALFSLFATGILEMRSGVGYDDEWNEOFWVGIGI 934
QY 933 SAHLFAVFOGLLKVLGADTNTFTVTSKANDREGDAELMYFKWTTLLIPPTTLLIINMWG 992
Db 935 SAHLFALFOGLLKVLGADTNTFTVTSKANDREGDAELMYFKWTTLLIPPTTLLIINMWG 993
QY 993 VVAGTSYAINSQSGWGLPGLKLEFAFWIVHLYPFLKGLMGQRNPTTIVVAVLLAS 1052
Db 994 VVAGTSYAINSQSGWGLPGLKLEFAFWIVHLYPFLKGLMGQRNPTTIVVAVLLAS 1053
QY 1053 IFSLLWRVDPFTTLAGPNQTQGCINC 1080
Db 1054 ILTLWVRVNPFFVAK--GGPVLEICGLNC 1080

RESULT 7
H84604
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84604
R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84604
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1088 <STO>
A:Cross-references: GB:AE002093; NID:g4417271; PIDN:AAD20396.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g21770
A:Map position: 2

Query Match 61.9%; Score 3579; DB 2; Length 1088;
Best Local Similarity 61.1%; Pred. No. 5.1e-254;
Matches 677; Conservative 156; Mismatches 199; Indels 76; Gaps 20;

QY 4 DADALKSRHGHGADVCOICADGLTLDGDVFTACDVCRFPVCRPCYEHKREGTQACIQ 63
Db 24 DTARISAEELSGQTKICRDEIELTDNGEFTIACNECAPCTPCYEYERREGNACPO 83
QY 64 CKTKYKRRHSGPAIRGEGDDTDADGSDFNYPASGTE-DOKOKIADRMRSWRMNT 117
```

Db 84 CGTRYKRIKSPRVEGDEEDD-DIDDLEHEFYGMD---PEHVTE-----AALYYMRLN 132
Qy 118 TG-GSGNVGHPKIDSGETGLSKYDSGEIPRGYVPSVNTSOMSGEIPGASPDHNMSP--- 173
Db 133 TGRGTDEYSH-----LYSAS--PGSEVPLLTVCDESDM--YSDRHAIIVPST 177
Qy 174 -TGNISRAPF-----PYNVHSPNPSREFS-GSIGNVANKERVDGWMKO-DKGA 220
Db 178 GLGNRVHVHPETDSFASHTPRMV-----PQKDLTVYGYGSVAWKRMEVWKKQIQEKLQ 232
Qy 221 IPTMNGTSIAPSEGRAADTIDASTEYNMEDAL-LNDETROPLSRKVPITASSKINPYRMI 279
Db 233 VVKNERVNDGGDGFIVDELD-----DPGLPMDEGRQPLSRKLPIRSRIRNIPYRMLI 285
Qy 280 VRLVVLVLIHYLHTNPVRNAYPLWLSVICEIWFALSWILDQFPKWFPIINRETYLDRL 339
Db 286 FCRLAILGLFFHYRILHPVNDFAFGLWLTSTVCEIWFASWILDQFPKWPYTERETYLDRL 345
Qy 340 ALRYDREGEPSOLAADVIFVSTVDPLEKPPITVANTVLSILAVDPVDKVSVCYVSDGAS 399
Db 346 SLRVEKSKGSELAPADVFTVDPLEKPPITVANTVLSILAVDPVEKVAICYVSDGAA 405
Qy 400 MLTEDALAESEFARKWVPFKYKIDIEPRAPFYPFOCKIDYLDKQVQPSFYKDRAMKRE 459
Db 406 MLTEALSYTAEFARKWVPFKYKIDIEPRAPFYPFOCKIDYLDKQVQPSFYKDRAMKRD 465
Qy 460 YEEFKIRINALVSKALKVPEEGWIMODGTPWPGNTRDPHGMIOVFLGHSGGLDTEGEL 519
Db 466 YEEFKVKNALVSKQKPEDGWTMQDGTWPGNVRDPHGMIOVFLGHSGVCDMDGNEI 525
Qy 520 PRLVYVSREKRPFGQHHKAGAMNALRVSAVLNTOYMLNDCDHYINNASKAVREAMCF 579
Db 526 PRLVYVSREKRPFGQHHKAGAMNLIIRVSAVLNAPYLLNVDCDHYINNASKAVREAMCF 585
Qy 580 LMDPNLGPOVCYVOPPOREFGIDRNDRYANRNTVFEDINLGLDGIQGVYVGTGCVFNR 639
Db 586 MMDPOSKKICYVOPPOREFGIDRNDRYANRNTVFEDINLGLDGIQGVYVGTGCVFNR 645
Qy 640 TAIYGEYEPPIKAKPGFLAS-----LCGGKKKASKSKRSDDKRSNKNHVDSSVPFN 692
Db 646 QALYGFADAPKKOPPGRTCNWPKWCCCLCGMKKTKGKVDNQKRPK---ETSKQIHA 702
Qy 693 LEDIEEGVEGAGFDEKSVLMSQMSLEKRFQGSAAVASTLMYEGGVQPSSTPESLKEA 752
Db 703 LEHIEEGQVNTAENNSPT--AQLKLEKKFGQSPVLVASTLLGLGVPSNVNPAASLLRES 760
Qy 753 IHVISCYGEDKSEWTEIGWIYGVSTEDILTFGRMHARGWRSVYCMKRPAPKGSAPINL 812
Db 761 IOVISCYGEERTWEGKEIGWIYGVSTEDILTFGRMHARGWRSVYCMKRPAPKGSAPINL 820
Qy 813 SDRNLQVLRWALGVSIEILFSRCHPLWYGGGRKLPLERFAYINTIYPLTSLPLLVYCIL 872
Db 821 SDRLQVLRWALGVSIEILFSRCHPLWYGGGRKLPLERFAYINTIYPLTSLPLLVYCIL 880
Qy 873 PAICLLTCKFTMPETSINLASWFTALFLSIFATGILEMRWSVGIDEMWNEQFWLTGGI 932
Db 881 PAICLLTCKFTMPETSINLASWFTALFLSIFATGILEMRWSVGIDEMWNEQFWLTGGI 940
Qy 933 SAHLFAVFGQLLKVLAGIDTFTVTSKANDEEGDAELMYKFWTLLPIPTTILINMVG 992
Db 941 SSSLFALPQGLLKVLAGVSTNFTVTSKAAD--DGEFSELYIFKWTSLPIPTTILINMVG 999
Qy 993 VVAGTSYAINSGXOSGWLFLGKLFAPWIVHLYPFLKGLMGRONRPTIIVHVAVLAS 1052
Db 1000 VIVGVSDAINGYDSWGLPFLGRLEFALWIVHLYPFLKGLMGRONRPTIIVHVAVLAS 1059
Qy 1053 IFSLLWRVDPFTTRLAGPNIOQTGCING 1080
Db 1060 ILLTLLWRVNFVSK-DGPVLEICGLDC 1086

RESULT 8

TI10797
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 16-Jul-1999 #sequence_re
C:Accession: TI10797
R:Pear, J.R.; Kawagoe, Y.; Schre
Proc. Natl. Acad. Sci. U.S.A. 93
A:Title: Higher plants contain h
A:Reference number: Z1152; MUI
A:Accession: TI10797
A:Status: preliminary; translate
A:Molecule type: mRNA
A:Residues: 1-974 <PEA>
A:Cross-references: EMBL:U58283;
A:Experimental source: strain AC
C:Genetics:
C:Function:
A:Description: involved in the s
C:Keywords: glycosyltransferase;
Query Match 60.2%;
Best Local Similarity 61.5%;
Matches 657; Conservative
Qy 18 VCQICADGLGTTLDGDVFTACD
Db 8 VCHTCGEHVGILNVGPEFVACH
Qy 78 RGEEDGTDADDGDSDFNYPASG
Db 59 -----DENLDD-----VEK
Qy 137 SKYDSGGEIPRGYVPSVTNSQMS
Db 97 STLDS-----EM 103
Qy 197 SGSTGNVANKERVGDGWMKQDR
Db 104 AEDNGSIWNRVSWKSKKRN
Qy 257 TROPLSKVPITASSKINPYRMV
Db 155 S-QPLSTIIIPKSRAPYRTV
Qy 317 LSWILDQPKWFPINRETYLDE
Db 214 FSWILDQPKWYVNNRETYIDE
Qy 377 LSTLAVDYPVKVSCYVSDDG
Db 274 LSTLAVDYPVKVSCYVSDDG
Qy 437 KIDYLDKQVPSFYKDRAMK
Db 334 KIDYLDKQVPSFYKERRANK
Qy 497 DHPGMIOVFLGHSGGLDTEG
Db 394 DHPGMIOVFLGHSGGLDTEG
Qy 557 YMLNLCDDHYINNASKAVREAM
Db 454 FILNLCDDHYINNASKAVREAM
Qy 617 INLRGLDGIQGVYVGTGCVF
Db 514 VNMKGLDGIQGVYVGTGCVF
Qy 677 KKSNNKHVDSSVPVNLNEDIE
Db 574 LYRDAKREELDAALFNREID

ytic chain celAI - upland cotton
n 16-Jul-1999 #text_change 21-Jul-2000
ost, W.E.; Delmer, D.P.; Stalker, D.M.
37-12842, 1996
gs of the bacterial cea genes encoding the cata
7296; PMID:8901635
m GB/EMBL/DDBJ
g1706955; PIDN:AAB37766.1; PID:g1706956
U-2; fiber
sis of cellulose
syltransferase
are 3477; DB 2; Length 974;
ad. No. 1.3e-246;
Mismatches 176; Indels 106; Gaps 15;
VCRPCYEHHERKEGTQACLOCKTKYKRRHSGPAI 77
ICKSCFEYDLKEGRKACLR-----GSPY- 58
OKTADMRSPRMNTGGSGNVG-HPKVDSGIGL 136
STMAA-----HLNKSQDVGIHARHS---SV 96
GASPDHMHMSPTGNISRRAPPYVNHSPNSREF 196
-----EM 103
TNGTSTAPSEGRAATDIDASTEYNMEDALLNDE 256
ATT-----KVERAEIPEQOQMEDKAPDA 154
LVLSIFLHYLTNPVRNAYPLWLSVICEIWF 316
LILGLFEHRYVTPNVDASAFGLWLTSTVICEIWF 213
YDREGEPSOLAADVIFVSTVDPLEKPPITVANTY 376
YEREGEPELAAVDFVSTVDPLEKPPITVANTY 273
PDALAESEFARKWVPFKYKIDIEPRAPFVEQ 436
FESLVETADFAKWWPFCKRFSIEPRAPFVFSQ 333
FKIRINALVSKALKVPEEGWIMODGTPWPGNNT 496
FKIRINALVAKAQKTDEGWTMQDGTWPGNNT 393
VYVSREKRPFGQHHKAGAMNALRVSAVLNTOY 556
VYVSREKRPFGQHHKAGAMNALRVSAVLNTOY 453
PNLGPQVCYVOPPOREFGIDRNDRYANRNTVF 616
POVGRDVCYVOPPOREFGIDRNDRYANRNTVF 513
YGYEPPITAKPGFLASLCCGGKKKASKSKRS 676
YGYGPPSPSPKSSSSSSSCSCCCKPKPKDPSE 573
AGFDD-EKSVLMSQMSLEKRFQGSAAVASTLME 735
-NYDEYKSLMISQTSFEKTFGLSVFIESTLE 627

QY 736 YGVPOQSPSPSLLKEAIIHVISCYEDKSEWGTGTEIGWISYVTEIDILTGFKMHARGWSV 795
DB 628 NGGVAESANPTSLKEAIIHVISCYEEKTAGKEIGWISYVTEIDILTGFKMHARGWSI 687
QY 796 YCMRPRAPKGSAPINLSDRNLQVLRWALGSVEILFRRHCPMLWGY-GGRLEKLEFAYI 854
DB 688 YCMPLRPAKGSAPINLSDRNLQVLRWALGSVEILFRRHCPMLWGYGGGGRLLQLRAYI 747
QY 855 NTIYPLSLPLLYCIIIPALCIIITGKFIPEISNLASIFIALFISIFATGILEMRWSG 914
DB 748 NTIVYFPTSLPIAYCSLPAICLTGKFIPTLSNLASVILFGLFISIIYAVLELRWSG 807
QY 915 VGDIDWNEQFQWVIGGSAHLFAVFCGLKLVLAGIDTNTVTSKANDSGDFAELMYMK 974
DB 808 VSIEDLRNEQFQWVIGGSAHLFAVFCGLKLVLAGIDTNTVTSKAAAD-DADFGELYIVK 866
QY 975 WTTLLIPTTILIIINWGVVAGTSYAINSGVQSWGPLEGKLFPAFWIVVILHYFLKGLMG 1034
DB 867 WTTLLIPTTILIIINWGVVAGTSYAINSGVQSWGPLEGKLFPAFWIVVILHYFLKGLMG 926
QY 1035 QNRPTTIVIMAVLLASIFSLWVRVDPFTRLAGPNI-QTC-GINC 1080
DB 927 QNRPTTIVIMAVLLASIFSLWVRVDPFTRLAGPNI-QTC-GINC 974

RESULT 9
T04870
cellulose synthase (EC 2.4.1.-) catalytic chain F28A21.190 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T04870
R:Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, February 1999
A:Reference number: 215387
A:Accession: T04870
A:Molecule type: DNA
A:Residues: 1-958 <BEV>
A:Cross-references: EMBL:AL035526
A:Experimental source: cultivar Columbia; BAC clone F28A21
C:Genetics:
A:Map position: 4
A:Introns: 60/1; 76/3; 94/1; 149/2; 240/2; 355/3; 401/3; 443/3; 514/3; 662/3; 780/3; 791/3
A:Note: F28A21.190
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 57.2%; Score 3302.5; DB 2; Length 958;
Best Local Similarity 59.5%; Pred. No. 8.2e-234;
Matches 638; Conservative 125; Mismatches 174; Indels 135; Gaps 17;

QY 18 VQICADGLGTTLDGVTACDVCRFPVCRPCYERKEGTQACLOCKTKYKRRHGPSAI 77
DB 8 ICNTGCEITGVKSGNEFFVACHCECFPCACLEVEFEKGRRCICLRCGNPY----- 58
QY 78 RGEEDDTDDGSDFNYPASTEDQKOKIADRMRSWMNTGGSGNVGHPKYDSGEIGLS 137
DB 59 -----DENVFDVE-----TKTSKQSID-----SGI--HAKHS---TVS 89
QY 138 KYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMMSPGTGNISRRAPFPYVNHSPNPREFS 197
DB 90 TIDS-----ELN 96
QY 198 GSGNVAVKERVGDKMKODKGAIPWTNGTSTAPSEGAATDIDASTEYNMEDALLNDET 257
DB 97 DEYGNPIWNRVSRWDRKDKSKKKKDKDPKATRAEQHEA---QIPQOHMEDTPPPNTES 153
QY 258 --RQPLSRKVPYASSKINPYRMVIVLRVLVLSIFLRLTNPNVRNAYPLWLLSVICBIWF 315
DB 154 GATDVLVSVIPRTPKITSYRIVIMRLIILALFPNRYTHPVDVSAAYGLWLTSVICBIWF 213
QY 316 ALSWILDQPKFPFPNRETYIDRLALRYDRCEPSQLAAVDIFVSTVDPLKEPPIVANT 375
DB 214 AVSWLDQPKFPNRETYIDRLSARFEREGEQSLAAVDFFVSTVDPLKEPPLITANT 273

QY 376 VLSILAVDYPVDKYSCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPAPPEFC 435
DB 274 VLSILAVDYPVDKYSCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPAPPEFC 333
QY 436 QKIDYLKDKVQSPFVKORAMKREYEEFKIRINALYSKALKVPEEGIMODGTPWPGNNT 495
DB 334 LKIDYLRDKVQSPFVKORAMKREYEEFKIRINALYSKALKVPEEGIMODGTPWPGNNT 393
QY 496 RDHFGMIOVFLGHSGGLDTFEGNELPLRVYVSRKRCFOHKKAGAMNALRVYSAVLWNG 555
DB 394 RDHFGMIOVFLGHSGGLDTFEGNELPLRVYVSRKRCFOHKKAGAMNALRVYSAVLWNG 453
QY 556 QYMLNLCDDHYINNSKAVREAMCFMDPNLGPOVYVQFPQRFQREDGIDRNDRYANRNTVFF 615
DB 454 PFILNLCDDHYINNSKAVREAMCFMDPNLGPOVYVQFPQRFQREDGIDRNDRYANRNTVFF 513
QY 616 DINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFL-----ASLCGKKKASKS 670
DB 514 DVNMRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFL-----SKPRILPOSSSSSCCLUTKK--KQ 568
QY 671 KKRSDKKKSNKHYDSSVPFNLIEDIEGVGAGFDD-EKSVLMSQMSLEKREFGQSAFV 729
DB 569 PQDPSEIYKDAKREELDAAIFNLGDL-----NYDEYDRSMLISQTSPEKTFGLSTVFI 622
QY 730 ASTLMEYGGVQSPSTPESLLKEAIIHVISCYEDKSEWGTGTEIGWISYVTEIDILTGFKMHA 789
DB 623 ESTLMENGSGVDSYNPSTLKEAIIHVISCYEEKTEGKEIGWISYVTEIDILTGFKMHC 682
QY 790 RGRSVYCMRPRAPKGSAPINLSDRNLQVLRWALGSVEILFRRHCPMLWGY-YGGRLEKFL 848
DB 683 RGRSVYCMRPRAPKGSAPINLSDRNLQVLRWALGSVEILFRRHCPMLWGY-YGGRLEKFL 742
QY 849 ERFVINTIYPLSLPLLYCIIIPALCIIITGKFIPEISNLASIFIALFISIFATGIL 908
DB 743 QRLAYINTIYPTTSLPLVAYCTLPALCIIITGKFIPEISNLASIFIALFISIFATGIL 790
QY 909 EMRWGSGVIDEWRNEQFVIGGISAHLFAVFCGLKLVLAGIDTNTVTSKANDSGDFA 968
DB 791 -----GVSIEDLRNEQFVIGGISAHLFAVFCGLKLVLAGIDTNTVTSKANDSGDFA 844
QY 969 ELYMKVWTTLLIPTTILIIINWGVVAGTSYAINSGVQSWGPLEGKLFPAFWIVVILHYYP 1028
DB 845 ELYMKVWTTLLIPTTILIIINWGVVAGTSYAINSGVQSWGPLEGKLFPAFWIVVILHYYP 904
QY 1029 LKGLMGRQNRPTTIVIMAVLLASIFSLWVRVDPFTRLAGPNIQTCGINC 1080
DB 905 LKGLMGRQNRPTTIVIMAVLLASIFSLWVRVDPFTRLAGPNIQTCGINC 953

RESULT 10
T10800
cellulose synthase (EC 2.4.1.-) catalytic chain celsa2 - upland cotton (fragment)
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10800
R:Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996
A:Title: Higher plants contain homologs of the bacterial celsa genes encoding the cat-
A:Reference number: #17152; MUID:97057296; PMID:8901635
A:Accession: T10800
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-685 <PEA>
A:Cross-references: EMBL:U58284; NID:gl706957; PIDN:AAB37767.1; PID:gl706958
A:Experimental source: strain Acala SU-2; fiber
C:Genetics:
A:Gene: celsa2
C:Function:
A:Description: involved in synthesis of cellulose
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 48.7%; Score 2816; DB 2; Length 685;
Best Local Similarity 74.0%; Pred. No. 2.5e-198;

Matches 509; Conservative 78; Mismatches 77; Indels 24; Gaps 6;

```
QY 414 RKWFEVKKYDIEPRAPFEYFCQIDYLDKQVQPSFYKDRRAMKREYEEFKIRINALVSK 473
D 1 RRPWFCKKHNVPRAPFEYNEKIDYLDKQVHPSFVKERRAMKREYEEFKIRINALVAK 60
QY 474 ALKVPBEGWIMODGTPWGNTRDHPGMIOVFLGHSGGLDTFEGNELPLVLYVVSREKRGF 533
D 61 AQKPEEGWIMODGTPWGNTRDHPGMIOVFLGHSGGLDTFEGNELPLVLYVVSREKRGF 120
QY 534 QHKKAGAMNALVRYSVAVLTNGQYMLNLDCHYINNSKAVREAMCFMDPDLNGLPQVQVQ 593
D 121 QHKKAGAMNALVRYSVAVLTNAPFLNLDCHYINNSKAVREAMCFMDPDLNGLPQVQVQ 180
QY 594 PQPQEDGIDRNDRYANRNVTFEDINRLGLDGIQGVVYVGTGCVNRTAIYIEPEPIAKK 653
D 181 PQPQEDGIDRNDRYANRNVTFEDINRLGLDGIQGVVYVGTGCVNRTAIYIEPEPIAKK 240
QY 654 -----PGFLASLCGGKKKASKSKRS-----SDKKK-----SNKHVDSVPVFN 693
D 241 PKMTCDCHPSCCCCCGGRKSKKKEKGLGLGLYKKKMMKGNVYKKGAPVFDL 300
QY 694 EDIEBVEGAGFDD-EKSVLMSQMSLEKRFQCSAAFAVASTLMEXGGVPSQSTPESLKEA 752
D 301 EBIIEGLE-GYEELKSTLMSQKNFKRFQSGSPVFIASITLMENGGLPEGTNSTSLIKEA 358
QY 753 HVISCGYEDKSEWTEIGWYGVSTEDILTGFKHMGWBSVYVCMKRPAPKGSAPNL 812
D 359 HVISCGYEDKSEWTEIGWYGVSTEDILTGFKHMGWBSVYVCMKRPAPKGSAPNL 418
QY 813 SDRNLQVLRWALGSVEILFSRHCPLMWYGGGLKFLERFAYINTIYPLTSLPLLVYCI 872
D 419 SDRNLQVLRWALGSVEILFSRHCPLMWYGGGLKFLERFAYINTIYPLTSLPLLVYCI 478
QY 873 PAICLLTGKFMPEISNLASTWFLALFSLFATGILEMNSGVGIDEMWNEQFWIGGI 932
D 479 PAICLLTGKFMPEISNLASTWFLALFSLFATGILEMNSGVGIDEMWNEQFWIGGI 538
QY 933 SAHLFAVQGLLKVLAGIDTFTVTSKANDEGDFAEILYMEFKWTTLLIPPTTILINNVG 992
D 539 SAHLFAVQGLLKVLAGIDTFTVTSKANDEGDFAEILYMEFKWTTLLIPPTTILINNVG 597
QY 993 VVAGTSYAINSGYSGWGLFQKLFPAFWVTVHLYPFLKGLMGRNRTPTIVVWAVLLAS 1052
D 598 VVAGTSYAINSGYSGWGLFQKLFPAFWVTVHLYPFLKGLMGRNRTPTIVVWAVLLAS 657
QY 1053 IFSLLWVRVDPPTFLAGPNTQTCGNC 1080
D 658 IFSLLWVRVDPPTFLAGPNTQTCGNC 685
```

RESULT 11

D86157
hypotheical protein F22d16.26 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence-revision 02-Mar-2001 #text_change 31-Dec-2001

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86157

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1181 <STO>

A:Cross-references: GB:AE005172; NID:g6056428; PIDN:AAF02892.1; GSPDB:GN00141

C:Genetics:
A:Map position: 1Query Match 40.1%;
Best Local Similarity 42.9%;
Matches 489; Conservative 1

```
QY 15 AGDVCOI--CADGLGTITLDGDV 40.1%
D 158 SGQICWLKGGDEKV-----V 42.9%
QY 73 GSPAIRCEEGDDTDADDGSDFN 489
D 210 DDP-----ETEEED----- 1
QY 133 EIGLSKYDSGEIPRGVYPSVTN 40.1%
D 228 QMGESKLDK---RLSVVVKSKFA 42.9%
QY 193 SREPSGSGIGNVANKERVVDGKK 489
D 269 -----GNAYWPK--DGYG 1
QY 253 LNDETROPRLSRKVPDIASSKINE 40.1%
D 297 --ERSKRPLTRKVSVAALISE 42.9%
QY 313 IWFALSWILDOFPKWFPIINRET 489
D 355 LWFALSWILDOFPKWFPIINRET 1
QY 368 PPIVANTVLSILAVDYPVDK 40.1%
D 415 PPIVANTVLSILAVDYPVEK 42.9%
QY 428 RAPEYFCQIDYLDKQVQPSFY 489
D 475 RNPEAYFGQKRNFLKRVRLD 1
QY 476 K-----V 40.1%
D 535 RAKKKQEMMMGNPNQETVIV 42.9%
QY 509 SGGLDTEGE----- 489
D 594 PNAEPVYGAEDAENLDTTD 1
QY 554 NGQYMLNLDCHYINNSKAVR 40.1%
D 654 NGPFILNLDCHYIYNSMALR 42.9%
QY 614 FFDINLRGLDGIQGVVYVGTG 489
D 713 FFDVSMRALDGLQGMVYVGTG 1
QY 674 SSDKKKNKHVDSSVPV---F 40.1%
D 768 PKAMKKKDDVE--SLPINGEY 42.9%
QY 731 S-TLMEXYGG----- 489
D 814 SIPVAEYQGRLIQDLOGKKN 1
QY 770 IGWYGSVTEIDILGFKMHAR 40.1%
D 874 VGWYGSVTEIDVVTGYRHN 42.9%
QY 830 LFSRHCPLMWYGGGLKFLER 489
D 934 FFSRNNAIF--ATPRMKFLOR 1
QY 890 LASTWFLALFSLFATGILEM 40.1%
D 934 LASTWFLALFSLFATGILEM 42.9%
```

re 2315; DB 2; Length 1181;
d. No. 3.4e-161;
Mismatches 256; Indels 214; Gaps 31;

```
VCRPPVCRPCYEHKREKGTQACLOCKTKYKRRH 72
D 3-CGFRICRDCYFCITSSGGGCGCKEPEYRDN 209
STEDQKOKIADMRMSWRMNTGSGNVGPKYDSG 132
D 227 FEDEAKPLP----- 227
SEIPGASPDHMHMSPGTGNISRRAPFPVYVNHSPN 192
D 268 SDF-----DHPRWL-----FETKGTGY 268
SAIPMTNCTSTAPSEGRAATDIDASTEYNMEDAL 252
D 296 SG-----NGYETPPEFG----- 296
IVLRVLVLSIFLHYRLTNPNVRNAYPLMLLSVICE 312
D 354 TALRLVGLGLFLTWVRHPNREAMWLMGMSSTCE 354
LALRYDREG--EP---SOLAAYDIFSVTVDPKE 367
D 414 LKERFENLNRNPKRSDLPGLDIDFVSTADPEKE 414
SDDGASMLTFDALAETSEFARKWVPFVKYKDIPE 427
D 474 SDDGALLTFEALAQATASFASTWTPFCRKHIEP 474
SAMKREYEEFKIRINAL-----VSKAL 475
D 534 RRVKREYDEEVRINSLPEAIRRRSDAYNVHEEL 534
TMDQGTMPG-----NNTDRHPGMIOVFLGH 508
D 593 MSDGSHWPGTWSSETDNGSDHAGIIQAMLAP 593
PRLVYVSREKRPFGQHHKKAGAMNALVRVSAVL 553
D 653 PMLVYVSREKRPFGYDHKKKAGAMNALVRTSAIS 653
LMDPNLGPQVCYVQFPQRFQDGDIDRNDRYANNTV 613
D 712 MLDRG-GDRICYVQFPQRFQDGDIDRNDRYANNTV 712
TAIYGYEPPIKAKKPGFLASLCGGKKKASKSKKR 673
D 767 TALYGFSPPRATEHHGWL---GRRKVKISLRR 767
EEGVEGAGFDDKESVLMSQMSLEKRFQCSAAFAVA 730
D 813 DDG-----DIESLL-----LPRKFGNSNFVA 813
---VQSTSPESLKEATHVISCYGEYDKSEWTE 769
D 873 SLAVPREPLDAATVAEALSVISCFEDKTEWGR 873
YCMKRPAPKGSAPINLSDRNLQVLRWALGSVEI 829
D 933 YCVTKRDAFRGTAPINLTDRUHOVLWRWATGSVEI 933
NTTIYPLTSLPLVYCIPLPAICLLTGKFMPEISN 889
D 991 NYGMYPTTSLFVLYCIPLPAISLFSQGFIVQSIDI 991
GIDEMWNEQFWIGGISAHULFAVFGQLLKLVLG 949
D 949 GIDEMWNEQFWIGGISAHULFAVFGQLLKLVLG 949
```



```
Db 180 -----NTHLTDQVD-----ENGQOR-----PMLPGGGG----- 203
QY 134 IGLSKYDSGEIPRGYVPSVTNSOMSGEIPGASPDHMMSPGTGNIS--RRAPFPVYVNSPNP 192
Db 204 -----KMERLSVMYKSNKSNAL-----MRSGTGFEDNRWLF----- 235
QY 193 SREFSGS--IGNVAMKERVGDGMKODKGAIPMTNGTSIAPSEGRAATDIDASTEYNMED 250
Db 236 --ETTCTYGVGNATWTKDGFSGKGDG-----DGDG--NGMEA 271
QY 251 ALLNDETROPLSRKVPYIASKKINPYRMVIVLRVLVLSIFLHLRLTNPNVNAVPLWLLSVI 310
Db 272 QDLMSRWPRLTKLIPAGVIFVYLLIFIRIVLFALEFLTWKRKHQNPDAVWLWGMVYV 331
QY 311 CELWFALSWILOQPKWFFINRETYLDRLALRYD--REGEP--SOLAADVDFVSVTDPL 365
Db 332 CELWFALSWILOQPKWFFINRETYLDRLALRYD--REGEP--SOLAADVDFVSVTDPL 365
QY 366 KEPPIVTANTVLSILAVDYPVKVSDGASMLTFDALAETSEFARKWVPFVKKYDI 425
Db 392 KEPLPLVTANTVLSILAAEYVPEKLSYVSDGGALLTFEAMAEASAFANIWVPECRKHAI 451
QY 426 EPRAPFEYFCOKIDYLDKRVQSFYKDRAMKREYEFKIRINALYS----- 472
Db 452 EPRNPDYSFLRKDPYKKNVSDFKVDRRRKREDFKVRVNSLDPDSIRRRSDAYHARE 511
QY 473 -----KALKVPEEGWIMQDGTWPG-----NNTRDHPGMIQVEL- 506
Db 512 EIKAMKWRQNRDDPEMPVKIPKATW--MADGTHWPGTWTLSASDHAKGDHAGIIQVMLK 570
QY 507 -----GHSGG---LDTEGNEPLRLVYVSRKRGFGQHHKKAGAMNALVRVSAVLTNG 555
Db 571 PPSDEPLHGVSEGLDLDVIRLPLLVYVSRKRGFGYDHNKKAGAMNALVRASATMSNG 630
QY 556 QYMLNLDCDHYLNNSKAVREACFELMDPNLQVQVQVQFPQFQDGDIDRNDYANRNTVFF 615
Db 631 PFLNLDCDHYLYNSALREGFCFMMDRG--GDRLCVQVQFPQFQDGDIDRNDYANRNTVFF 689
QY 616 DINLRGLDGIQGVYVGTGCVFNRTAIYGEPPKIRAKKPGFGLASLGGKKKASKKRS 675
Db 690 DYNMRALDGLMGPVYVGTGCLFRRIALYGFNPPRSKDFSPSCWCC-----FPRSKK-- 742
QY 676 DKKSNNKHVDSSVPFVNLDEIEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFFVAS--TLM 734
Db 743 -----NIPENRALRMSDYDDEE---MNLSLVPKFKGNSFLIDSIPVA 783
QY 735 EYGVGPQSSTP-----ESILKEAIIHVISGVEDKSEWTEIGWIYG 775
Db 784 EQQRLADHPAVKNGRPPGALTIPRELLDASTVAEIAVISCWYEDKTEWGRSGRIWYG 843
QY 776 SVTEDILTGFKNHARGWSVYCMKRPAPKGSAPINLSRDLNQVLRWALSGVSEILFSSRC 835
Db 844 SVTEDVVTGYRMHNRGWSVYCVTKRADPRGTAPINLTDRLHQVLRWATGSEIIFSRNN 903
QY 836 PLWYGGGLKFLERFAYINTIYPLTSLPLVYCILPAICLLTGKFMPEISNLSAIWF 895
Db 904 ALL--ASSKMKILQRIAYLWNGIYPTSLIYVCFPLPALSFGQFIYQTLNVTFLVYL 961
QY 896 IALFLSIATGILEMRWSGVGIDENWRNEQFWVIGGISAHFLFAVFOGLLKLVLGIDTNT 955
Db 962 LIISITLCALLLEIKWSGISLEWRNEQFWLIGTSHAHLAAVLOGLLKVAVGVEIST 1021
QY 956 VTSKA--NDEEGFALYMKFWTLLIPPTTILINMVGVAGTSYAINSGYQSMGPLFG 1013
Db 1022 LFSKSGGDIDDEFADLYLVKVTSLMIPPTIIMVNLIAVGFSTIYSVVPQWSKLIG 1081
QY 1014 KLFFAFWLVHLVPLKGLMGQNKPTPTIYVWAVLLASIFSLLLWVRDP 1063
Db 1082 GVFFSFWLAHLVYPAKGLMGRGRRTPTIYVWVSGVLVAITISLLWVAINP 1131
RESULT 14
DB4741
```

```
probable cellulose synthase [imp
C:Species: Arabidopsis thaliana
C:Date: 02-Feb-2001 #sequence_re
C:Accession: D84741
R:Bin, X.; Kaul, S.; Rounsley, S
M.; Koo, H.; Moffat, K.S.; Croni
euss, D.; Nierman, W.C.; White,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis o
A:Reference number: A84420; MUI
A:Accession: D84741
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1036 <STO>
A:Cross-references: GB:AE002093;
C:Genetics:
A:Gene: At2g33100
A:Map position: 2
```

```
Query Match 38.3%;
Best Local Similarity 43.5%;
Matches 481; Conservative 1
```

```
QY 67 KYKRRGSPAIRGEGDDTDAD
Db 25 KFGRTSSGRIVLSLSD--DMD
QY 109 DRMSWRMTTGGSGNGVHPKYD
Db 84 D-----ANRGGGGGGPKMGN
QY 169 HMMSPGTGNIS--RRAPFPVYVNS
Db 117 -LRSQTGDDHNRWLF-----
QY 223 MTNGTSIAPSEGRAATDIDAST
Db 160 -----
QY 283 LVVLSTFLHYRLTNPNVNAVPL
Db 190 LVLVFLLWMRITNPNEADAWL
QY 343 YDREGF-----SOLAADV
Db 247 HDKFEQPSNPNTGRSDLPQVD
QY 395 DGASMLTFDALAETSEFARKW
Db 307 DDGAILTFEAMAEAVRFAEY
QY 455 AMKREYEFKIRINALVSKALK
Db 367 WKREDFEKVRLNGLPQIKK
QY 484 MDGTPWPG-----NNTRD
Db 426 MADGTHWPGTPEPKPDHSGD
QY 523 VVVSREKRGFGQHHKKAGAMNA
Db 486 AVVSREKRGFGQHHKKAGAMNG
QY 583 PNLGPOCVVQFPQFQDGDIDR
Db 546 RG-GDRICVQFQFQEGIDPS
QY 643 YGYEPPKAKPGFLASLCGGK
Db 605 YGFNPFRANEYSGVF-----GQ
QY 703 AGFDDDEKSVLMSQMSLEKRFQ
```

```
] - Arabidopsis thaliana
e-ear cross)
n 02-Feb-2001 #text_change 02-Feb-2001
Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
isen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
onosome 2 of the plant Arabidopsis thaliana.
3487; PMID:10617197
```

```
g2924781; PIDN:AAC04910.1; GSPDB:GN00139
re 2212.5; DB 2; Length 1036;
d. No. 9.2e-154;
Mismatches 266; Indels 217; Gaps 30;
-----NYPASGTEDQKOKIA 108
SGQNDYINTVLMPPTPDNPAGSGSTSESKG 83
LSKYDSGEIPRGYVPSVTNSOMSGEIPGASPDH 168
LS-----VMKSNKSM----- 116
EFSG--SIGNAWKER---VDGWMKQDKGAIP 222
ESKGYGIGNAFWSEEDDYDGVGSKSD----- 159
DALLNDETROPLSRKVPYIASKKINPYRMVIVLR 282
: : : : : : : : : : : : : : : : : :
---FLDKPKPLTRKVIQIPAKILSPYRLIIVIR 189
ICEIWFALSWILOQPKWFFINRETYLDRLALR 342
: : : : : : : : : : : : : : : : : :
VCEIWFALSWILOQPKWFFINRETYLDRLALR 342
VDPLEPPVTANTVLSILAVDYPVKVSCYVS 394
: : : : : : : : : : : : : : : : : :
ADPEKEPLVTANTVLSILAVDYPVIEKLSAYIS 306
KYDIEPRAPFEYFCQKIDYLDKRVQSFYKDRR 454
KHDIEPRNPDYSFISIKKDKPTKKNKQDFVKDR 366
-----VPEEG-----WI 483
NMREELKEKRIAREKNGVLPDGPVVEVWATW- 425
QVF-----LG---HSGGLDTEGNEEL--PRL 522
: : : : : : : : : : : : : : : : : :
QIMSKVLPDLEPVVGGPNEGALDFTGIDIRVPMF 485
AVLTNGOYMLNLDCDHYLNNSKAVREACFELMD 582
: : : : : : : : : : : : : : : : : :
AALSNGAFILNLDCDHYLYNSALREGFCFMMDR 545
RNTVFFDINLRGLDGIQGVYVGTGCVFNRTAI 642
: : : : : : : : : : : : : : : : : :
HNTVFDGNRRLDGLQGVYVGTGCMFRYAL 604
SKRSDDKKSKNNKHVDSSVPFVNLDEIEGVEG 702
: : : : : : : : : : : : : : : : : :
MHWRTQSQASQTSQA-----SDLESQTQ 650
AS-TLMEYGVGP-----QSTPESLL---- 749
: : : : : : : : : : : : : : : : : :
```


Db 651 LNDPDP-----LGLPKKFGNSTMTDTIPVAEYOGRLADHMSVKNRPPGALLPRP 703
QY 750 -----KEAIHVISCGYEDKSWGTEIGWYGSVTEIDLTPKMHARGWSRYCMKPRP 802
Db 704 PLDAPTVAEIAVISWYEDTNGRIGWYGSVTEIDVTVGYRMEHNRGWSRYCITKRD 763
QY 803 AFKGSAPINLSRLNOVLWALGSVELFSRCHPCPLAWGYGGYGRLLKFLERFAYINTTYPLT 862
Db 764 AFRGTAPINTDRLHOVLWATSGVEIFFSKNNMFE--ATRRUKFLQRYAVLNVGVIYPT 821
QY 863 SLPLLYCILLPAICLTGKFMPEISINLASINFIALFLGIFAT-----GILEMRWSGVGD 918
Db 822 SIFLWYCFPLALCSGKFIQVSL-----DIHFLSYLLCITVTTLTSLLEVKWSGIGLE 877
QY 919 EWRNEQFWWIGISAHFLFAVFOGLLKVLAGIDTNTVTYSKANDERGD--FAELYMFKWT 976
Db 878 EWRNEQFWWIGISAHFLFAVFOGLLKVLAGIDTNTVTYSKANDERGD--FAELYMFKWT 937
QY 977 TLLIPPTTILINNMGVWAGTSVAINSGYSGWGLFGKLFFAFWVIVHLYPFLKGLMGRO 1036
Db 938 GLFIMPITLIIINLVAVIGASRTIYSVIPQWCKLGGIFFSLVWVTHMYPFAKGLMGRR 997
QY 1037 NRTPTIVWVAVILLASIFSLWVRDP 1063
Db 998 GKVPTIVYVNSGLVSITVSLWITISP 1024

RESULT 15
C86446
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: C86446
R/Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, K.;
Nature 408, 815-820, 2000
C/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maifi, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: C86446
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-979 <STO>
A/Cross-references: GB:AE005172; NID:g10801364; PIDN:AAG23436.1; GSPDB:GN00141
A/Genetics:
A/Map position: 1

Query Match 37.4%; Score 2163.5; DB 2; Length 979;
Best Local Similarity 44.3%; Pred. No. 3.3e-150;
Matches 450; Conservative 173; Mismatches 245; Indels 147; Gaps 24;

QY 142 GEIPRGVVPVSTNSGSGEIPGASPDHMHMSPTGNTSRAPPFPVYVNHSPNSRPFSGSIC 201
Db 4 GESPLRH-PRISHVNSGSGFSGSSYKNY-----LVQIPPTP--DNNPGPA-----SUS 50

QY 202 NVAMKRVGDKMKQDKGAIPMTNGTSIAPSEGR-----AATDIDASTEYNMEDALLNDET 257
Db 51 IVL-----LEIDSQESVPSVSGDIVSGSGCKDNEPDLTQVRINVNGEEEDDTLLSKI 103

QY 258 RQPLSRKVPITASSKINPYRVIVLWVLSIFLHYRLTNPNRYNAYPIWLLSVCEIWFAL 317
Db 104 SYSLTRVYKISPIIALYRILIVRVVSLALFLWRIRPNPNKALWLLSVLCEIWFALF 163

QY 318 SWILDQPKWFPNRTYDLRLALRY-----DREGESQLAADVIEVSTVDPLKEPIVT 372
Db 164 SWILDQPKWFPNRTYDLRLALRY-----DREGESQLAADVIEVSTVDPLKEPIVT 223

QY 373 ANTVLISILAVDVPVVKVSYSDDGASMLTFDALETSEFARKWVPFVKYDIEPRAPEF 432

Search completed: February 19, 2003, 16:42:41

Job time : 59 secs

Db 224 ANTILISUSVDPVEKLSVYISDDGSLVTEFAETAASAFKIVWPCRKHKIEPRAPES 283
QY 433 YFCQKIDYLDKVKVOPSFVKDRAMKREYEEFKIRINALV-----SK----- 473
Db 284 YFGLKRDYKDKVRHDFVRRRYVKRAYDEFKVRNALPHSIRRSADFNSKEEIKALEK 343
QY 474 -----ALKVPEBGIMQDGTWPQG-----NNTRDHPGMIQVFL- 506
Db 344 WKHKVKEEDQIKRPPALVAPKATW-MSDGTWHPGTWAVSGPHHSGDHASVIOVLDD 402
QY 507 -----GHSGLDTEGNE--LPRLVVVSREKRPGEFOHKKKAGAMNALVRSVAVLT 553
Db 403 PPGDEPVEGKGEGALDLEGVDIRLPMVYVSREKRPGEFOHKKKAGAMNALVRSVAVLT 462
QY 554 NGQYMLNLDCHYINNSKAVREACMFLMDPNLGPQVCYVQPPQRFQDIDRNDRYANRTV 613
Db 463 NGPFTLNLDCHYVYNSRAFRDGCIFMMDHD-GDRVSYVQPPQRFQDIDRNDRYANRTV 521
QY 614 PFDINRLGIDGLOGPVYVGTGCVENRTAIYGEPP---IRAKKPGFLASLGGGKKASKS 670
Db 522 PFDINRLGIDGLOGPVYVGTGCVENRTAIYGEPP---IRAKKPGFLASLGGGKKASKS 574
QY 671 KKRSSDKKSKNKHVDSSVPVENLEDIERGVGAGDFDDEKSVLMSQMSLEKRFQSOAAFA 730
Db 575 KKRSPATVAS-----EPEYTTDEEDRFDIGL-IRKQFGSSMLVN 613
QY 731 ST-LMEYGGVQSSST-----PESL-----LKEAIHVISCGYEDKSEWTEI 770
Db 614 SVKVAEFGRRPLATVHSSRLGRPPGSLTGSRKPLDFATVNEAVNVISWYEDKTEWGFNV 673
QY 771 GWYGSVTEIDLTPKMHARGWSRYCMKPRKPAKGSAPINLSRLNOVLWALGSVEL 830
Db 674 GWYGSVTEIDLTPKMHARGWSRYCMKPRKPAKGSAPINLSRLNOVLWALGSVEL 733
QY 831 FSRHCPLWYGGGRLLKFLERFAYINTTYPLTSLPLVYLCILPAICLTGKFMPEISNL 890
Db 734 FSRNNAIF--AGPKLLQLQRTAYLNVGIYPTSFILTYCFLPPLSLFSGHFVETLTGS 791
QY 891 ASIWFIALPLSIFATGILEMRWSGVGIDENWRNQFWWIGISAHFLFAVFOGLLKVLAGI 950
Db 792 FLIYLLIITLSLCLGLAVLEVKWSGISLEEMWRNEQFWLIGTSAHLVAVLQGIKVIAGV 851
QY 951 DTNFTVTSKA---NDEEGDFAELYMFKWTLLIPPTTILINNMGVWAGTSVAINSGY 1006
Db 852 EISFTLTSKSSTGGDDEDEDFADLYLFWKALMIPPLTIIILNIVAILFAVCRVTFSANP 911
QY 1007 SWGPLFGKLFPAFWVIVHLYPFLKGLMGRQNRTPPTIVWVAVILLASIFSLWVRV 1061
Db 912 QWSNLLGCTFFASWVLLHMYPFKGLMGRGKTPVTVVYVWSGLIAICLSLIYITI 966

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: February 19, 2003, 16:34:11 ; Search time 34 Seconds
(without alignments)

1317.483 Million cell updates/sec

Title: US-09-900-237-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHGAGDVQ.....VDFFTTRLAGPNIQTCGINC 1080

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	264	4.6	874	1 BCSA_SALTI	Q8z291 salmonella
2	264	4.6	874	1 BCSA_SALTY	Q93in2 salmonella
3	259.5	4.5	1596	1 ACS2_ACEXY	Q59167 acetobacter
4	258	4.5	872	1 BCSA_ECO57	Q8x517 escherichia
5	257	4.4	872	1 BCSA_ECOLI	P37653 escherichia
6	248	4.3	739	1 BCSA_PSEFL	P58931 pseudomonas
7	238.5	4.1	754	1 BCAL_ACEXY	P19449 acetobacter
8	238.5	4.1	1518	1 BCAL_ACEXY	Q9rb12 acetobacter
9	237.5	4.1	1518	1 BCSA_ACEXY	Q9wx75 acetobacter
10	237.5	4.1	756	1 BCAL_ACEXY	Q82859 acetobacter
11	230	4.0	729	1 BCSA_XANAC	P58932 xanthomonas
12	207	3.6	1550	1 ACS1_ACEXY	P21877 acetobacter
13	202.5	3.5	745	1 BCSA_ACEXY	Q9wx61 acetobacter
14	119	2.1	1433	1 SUBF_BACSU	P16397 bacillus su
15	116.5	2.0	441	1 YCDQ_ECOLI	P75905 escherichia
16	112.5	1.9	322	1 NU1M_STRPU	P15548 strongyloce
17	112	1.9	1029	1 RIP3_RAT	Q9ere6 bacillus norv
18	111	1.9	414	1 YOAB_BACSU	Q34864 bacillus su
19	110.5	1.9	1154	1 KDGD_MESAU	Q64398 mesocricetu
20	109.5	1.9	1024	1 RIP3_MOUSE	P97434 mus musculu
21	108.5	1.9	1262	1 TPP2_MOUSE	Q64514 mus musculu
22	108	1.9	1167	1 CLAA_BACTU	P56956 bacillus th
23	108	1.9	1894	1 ME2L_SCHPO	O13967 schizosacch
24	107.5	1.9	775	1 TCP0_HSV11	P08393 herpes simp
25	107.5	1.9	920	1 DMF1_SCHPO	P78953 schizosacch
26	107.5	1.9	1671	1 DPOL_PYRKO	P77933 pyrococcus
27	107.5	1.9	3033	1 POLG_HCVJ6	P26660 h genome po
28	107	1.9	1249	1 TPP2_RAT	Q64560 rattus norv
29	106	1.8	357	1 G6PT_CANFA	O19133 canis famil
30	105.5	1.8	662	1 CYOB_BUCAL	P57543 buchnera ap
31	105.5	1.8	886	1 YKRL_CAEEL	P34307 caenorhabdl
32	105	1.8	357	1 G6PT_HUMAN	P35575 homo sapien
33	104.5	1.8	541	1 COX1_PODAN	P20681 podospora a

RESULT 1

BCSA_SALTI	STANDARD;	PRT;	874 AA.
ID	BCSA_SALTI		
AC	Q8Z291;		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DE	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).		
GN	BCSA OR STV4181.		
OS	Salmonella typhi.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Salmonella.		
OX	NCBI_TaxID=601;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CT18;		
RX	MEDLINE=21534947; PubMed=11677608;		
RA	Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,		
RA	Churcher C., Muncall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,		
RA	Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,		
RA	Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,		
RA	Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,		
RA	Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,		
RA	Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,		
RA	Whitehead S., Barrell B.G.;		
RT	"Complete genome sequence of a multiple drug resistant Salmonella		
RT	enterica serovar Typhi CT18.";		
RL	Nature 413:848-852(2001).		
CC	!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes		
CC	uridine 5'-diphosphate glucose to cellulose, which is produced as		
CC	an extracellular component for mechanical and chemical protection		
CC	at the onset of the stationary phase, when the cells exhibit		
CC	multicellular behavior (rdar morphotype). Co-expression of		
CC	cellulose and thin aggregative fimbriae leads to a hydrophobic		
CC	network with tightly packed cells embedded in a highly inert		
CC	matrix (By similarity).		
CC	!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4-beta-D-glucosyl))(N) = UDP		
CC	+ ((1,4-beta-D-glucosyl))(N+1).		
CC	!- COFACTOR: Magnesium (By similarity).		
CC	!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid		
CC	(C-di-GMP) (By similarity).		
CC	!- PATHWAY: Bacterial cellulose biosynthesis.		
CC	!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane		
CC	(Potential).		
CC	!- DOMAIN: There are two conserved domains in the globular part of		
CC	the protein: the N-terminal domain (domain A) contains		
CC	the conserved DXD motif and is possibly involved in catalysis and		
CC	substrate binding. The C-terminal domain (domain B) contains the		
CC	QXXRW motif and is present only in processive glycosyl		
CC	transferases. It could be involved in the processive function of		
CC	the enzyme, possibly required for holding the growing glycan chain		
CC	in the active site.		
CC	!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		

CC -!- PATHWAY: Bacterial cellulose biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- DOMAIN: There are two conserved domains in the globular part of
 CC the protein: the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC OXRRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ315770; CAC44015.1; -;
 CC DR EMBL: AE008667; AAL22479.1; -;
 CC DR EMBL: AJ315148; CAC86199.1; -;
 CC DR StyGene; SG2222; bcsA.
 CC InterPro; IPR005150; Cellulose_synt.
 CC InterPro; IPR001173; Glycos_transf_2.
 CC Pfam; PF03552; Cellulose_synt; 1.
 CC Pfam; PF00535; Glycos_transf_2; 1.
 CC KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
 CC Transmembrane; Inner membrane; Complete proteome.
 CC FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.
 CC FT DOMAIN 441 501 CATALYTIC SUBDOMAIN B.
 CC FT TRANSMEM 30 50 POTENTIAL.
 CC FT TRANSMEM 151 171 POTENTIAL.
 CC FT TRANSMEM 173 193 POTENTIAL.
 CC FT TRANSMEM 230 250 POTENTIAL.
 CC FT TRANSMEM 525 545 POTENTIAL.
 CC FT TRANSMEM 547 567 POTENTIAL.
 CC FT TRANSMEM 592 612 POTENTIAL.
 CC FT TRANSMEM 634 654 POTENTIAL.
 CC FT TRANSMEM 668 688 POTENTIAL.
 CC FT TRANSMEM 833 853 POTENTIAL.
 CC FT ACT_SITE 313 313 POTENTIAL.
 CC FT ACT_SITE 457 457 POTENTIAL.
 CC FT SITE 360 360 SUBSTRATE BINDING (POTENTIAL).
 CC FT SITE 362 362 SUBSTRATE BINDING (POTENTIAL).
 CC SQ SEQUENCE 874 AA; 100043 MW; 4C9421B58606310A CRC64;

Query Match 4.68; Score 264; DB 1; Length 874;
 Best Local Similarity 19.58; Pred. No. 1.2e-10;
 Matches 159; Conservative 91; Mismatches 220; Indels 346; Gaps 30;

QY 261 LSRKVPFIASSINPRVIVLRLVLSIFLHYRLT-----NPNRYAPLWLISVCEIW 314
 Db 189 LVRRNPGFSAI-----MLVLSLVSCRYINWRTYSTLNDDPVSIVCGILLIFAEYAW 244
 QY 315 FALSILDOFPKPFPPINRETYDLRALRYDRGEPGSOAAVDIFSVDPKPEPIVTAN 374
 Db 245 IVL--VLGYFQVWPLNRQP-----VPLPKEMSQWPTVDIFVPTYN---EDLNVRKN 291
 QY 375 TVLSLIVADYVPDKVSCVSDGASMLTFDALAETSEFAKWKVPFVKKYDIEPRAPEFF 434
 Db 292 TIYASLGIDWPXKDLNIWLDG-----
 QY 435 CQKIDYLDKQVPSFVKDRAMKREYEEFKIRINALYSKALKVPEEGWIMQDGTWPGNN 494
 Db 315 -----GRE 317
 QY 495 FPDHPCMTQVFLHSGGLDTGNEPLRLVYVSREKRPFGQHHKKAGAMNALVRYSAVLTN 554
 Db 318 SFRH-----FARHVG-----VHYIARTT-----HEHAKAGNINNALKHA---K 352

QY 555 GQVNLNDCDHY INNSKAVREAMCFMLDPNLGPQVCYVOPPORF-----DGIDRN-----DRY 607
 Db 353 GEFAIFDCDCHVTRPSFLQMTMGWFLKE-----KQLAMMQTPHHFFSPDFERNLGRFRT 408
 QY 608 ANRNTVFDDINLRGLDGIQGPVYVGTGVNRTAIYGYEPIKAKKPGFLASLCGGKKKA 667
 Db 409 PNEGTFLYGLVQDGNMDWDATFFCGCAVIRR-----RP-----442
 QY 668 SKSKRRSSDKKSNKHVDSSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEKRFQGSAA 727
 Db 443 -----442
 QY 728 FVASTLMEYGVGPOSSTPESLLKEAIIHVISCGYEDKSEMGTEIGWYGSVTEIDILGFKM 787
 Db 443 -----LDEIGI-----AVE-----TWTEAHTSLRL 464
 QY 788 HARGWRSYCMKPKPAFKSAPINLSDRNLQVLRWALGSVEILFSRHCPGLWYGGRLKF 847
 Db 465 HRRGITYAYM--RIPQAGLATSELSAHIGQIRWARGWVOI--FRIDNPL---FGGLKL 518
 QY 848 LERFAYINTTIYPLTSLPLVYLCILPAICLLTGKFTIMPEISNLASIWFIALFISIFA-TG 906
 Db 519 AQLCYLNAHFHFLSGIPRLIFLTAFLALLHAYI-----IYAPALMIALEVIPH 569
 QY 907 ILEMWSGVGIDEWNRNEQFNVIGGISAHLEFAVFGQLLKVLAGIDTN---FTVTSKANDE 963
 Db 570 MVHASLNSKIQGYRH-SFW--SEIYETVLAWYIAPPTLVALINPHKGFNVTAK-----622
 QY 964 EGDEAELMYMFKWTTLLIPPTTILINMGVYVAGTSYAINSGYOSWGLPFGKLFPAFWIV 1023
 Db 623 -GGIVEEKYVDW-VISRPYIFLVLLNLGVAAGV-----WRYYYGP-----661
 QY 1024 HLYFPLKGLMGRQNRPTTIVVWVALLASIFSLWV 1059
 Db 662 -----ENETLVIV-----SLVWV 675

RESULT 3
 ACS2_ACEXY STANDARD; PRT; 1596 AA.
 ID ACS2_ACEXY
 AC Q59167;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cellulose synthase 2 [Includes: Cellulose synthase catalytic subunit
 DE (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose
 DE synthase 2 regulatory domain)].
 GN ACSAII.
 OS Acetobacter xylinus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconacetobacter.
 OX NCBI_TaxID=28448;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 23769;
 RX MEDLINE=95394846; PubMed=7665515;
 RA Saxena I.M., Brown R.M. Jr.;
 RT "Identification of a second cellulose synthase gene (acsAII) in
 RT Acetobacter xylinum.";
 RL J. Bacteriol. 177:5276-5283(1995).
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4-beta-D-glucosyl))(N) = UDP
 CC + ((1,4-beta-D-glucosyl))(N+1).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Potential).
 CC -!- DOMAIN: There are two conserved domains in the globular part of
 CC the catalytic subunit: the N-terminal domain (domain A) contains
 CC the conserved DXD motif and is possibly involved in catalysis and
 CC substrate binding. The C-terminal domain (domain B) contains the
 CC OXRRW motif and is present only in processive glycosyl
 CC transferases. It could be involved in the processivity function of
 CC the enzyme, possibly required for holding the growing glycan chain
 CC in the active site.
 CC -!- MISCELLANEOUS: It is not essential for cellulose production in

D	b	188		--RRKAFHDFAVRAGA-----GYIIR-----	206
Q	y	512	LDTGNEGLPRLVYVSREKRPGFQHKKAGAMNALVRYSVLVTNGYYMLNDCDHYINNSK	571	
D	b	207	--AENN-----HAKAGNLNHALAV-----TDSPPFAVFIEDCDHVPTRGF	242	
Q	y	572	AVREAMCFLMDPNLGNPOVCYVQFPORFDGIDRNDRYNRNTVPFDINLRGLDGTGGPVVV	631	
D	b	243	LRTIGMMADPNL-----ALLQTPIHF-----	265	
Q	y	632	GTCGVFNRTAIYGEPPIKAKKPGFASLCCGKKKASKSKRSDDKKSNKHVDSSVPVF	691	
D	b	266	-----YAPDP-----FORNLAGG-----MHVPPEGNMFF	288	
Q	y	692	NLEDIEBEGVGAGFDEKSVLMASQMSLEKREGQSAAFVASTLMEYGGVQPSSTPESLKE	751	
D	b	289	-----YGLVQDGND-----FWDATFFCGSCAI-----IRRE	314	
Q	y	752	AIHVISCGYEDKSEWGTEIGWYGSVEDILTGFKMHHARGWSRYVCMKRPKFGKSAPIN	811	
D	b	315	AVMGIG-----GFATETVTDATHAKMQRWGWTAYL--REPLAAGLATER	359	
Q	y	812	LSDRLNQWRWALGSVEILFSRHCPLVWYCYGGRCLKFLEREAYINTTIYPLTSLPLLVCYI	871	
D	b	360	LIUHGQRVWRWAGMTQIMRLNDPML-----GAGLRWEQRCLYLSAMGHFLPAIPRLTFIV	415	
Q	y	872	LPAICLLTGKFMPEISNIASIWFIA-LFLSIFATGILEMWSGVGIDEWMNRNQFW-VI	929	
D	b	416	SPLAFLFLGONITAAASPVAISVALPHIFSHVITLSRIEGRW-----RYSFWSEI	465	
Q	y	930	GGISAHLFAVFOGLLKVLVLAGIDTNFTVTISKANDESGDAELYMKFKWTLTLPPTTI---	985	
D	b	466	YETSLLALFLVRITIVILLQPHKGKNVT-----DKGGLLARGY-FDWDA--VYPNVILAGV	518	
Q	y	986	-----LINMVGVV-----AGTSYAINSQSYQSOGPLFGKLFPFAPVTVIHLVPLKGL-M	1033	
D	b	519	LCAALLRGVGIWQPHDLALQSETLNT-----LWVVISLIIVLASIAV	563	
Q	y	1034	GRONR 1038		
D	b	564	GRETR 568		
 RESULT 9					
B	C	IDA5_ACETY	STANDARD;	PRT; 1518 AA.	
D	C	AC Q9WX75;			
T	C	15-JUN-2002 (Rel. 41, Created)			
D	T	15-JUN-2002 (Rel. 41, Last sequence update)			
D	T	15-JUN-2002 (Rel. 41, Last annotation update)			
D	E	Putative cellulose synthase 3 [includes: Cellulose synthase catalytic subunit (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose synthase 3 regulatory subunit)].			
D	E	BCSABII-B.			
G	S	Acetobacter xylinus.			
O	C	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;			
O	C	Glucanacetobacter.			
O	X	NCBI_TaxId=284448;			
R	N	[[]]			
R	P	SEQUENCE FROM N.A.			
R	K	STRAIN=JCM 7664 / IFO 13693;			
R	C	PubMed=10382968;			
R	A	Uneda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M., Inoue Y.			
R	A	RNA			
R	T	"Cloning of cellulose synthase genes from Acetobacter xylinum JCM 7664: implication of a novel set of cellulose synthase genes.";			
R	T	DNA Res 6:109-115(1999).			
E	L	- CATALYTIC ACTIVITY: UDP-glucose + {(1,4-beta-D-glucosyl)}(N) -> UDP-glucose + {(1,4-beta-D-glucosyl)}(N+1).			
C	C	- PATHWAY: Bacterial cellulose biosynthesis.			
C	C	- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).			
C	C	- DOMAIN: There are two conserved domains in the globular part of			

the catalytic subunit: the N-terminal domain (domain A) contains the conserved DXD motif and is possibly involved in catalysis and substrate binding. The C-terminal domain (domain B) contains the QXNRW motif and is present only in processive glycosyl transferases. It could be involved in the processivity function of the enzyme, possibly required for holding the growing glycan chain in the active site.

-!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

DR	EMBL: AB010645; BAA31463.1; -	
DR	InterPro: IPR003919; CellSynth_A.	
DR	InterPro: IPR001173; Glycos_transf_2.	
DR	Pfam: PF00535; Glycos_transf_2; 1.	
DR	PRINTS: PR01439; CELLSNTHASE.	
DR	Cellulose biosynthesis; Transferase; Glycosyltransferase;	
KW	Transmembrane; Inner membrane.	
FT	DOMAIN 147 242	CATALYTIC SUBDOMAIN A.
FT	DOMAIN 319 379	CATALYTIC SUBDOMAIN B.
FT	TRANSMEM 27 47	POTENTIAL.
FT	TRANSMEM 49 69	POTENTIAL.
FT	TRANSMEM 106 126	POTENTIAL.
FT	TRANSMEM 167 187	POTENTIAL.
FT	TRANSMEM 409 429	POTENTIAL.
FT	TRANSMEM 432 452	POTENTIAL.
FT	TRANSMEM 470 490	POTENTIAL.
FT	TRANSMEM 517 537	POTENTIAL.
FT	TRANSMEM 551 571	POTENTIAL.
FT	ACT_SITE 189 189	POTENTIAL.
FT	ACT_SITE 335 335	POTENTIAL.
FT	SITE 238 238	SUBSTRATE BINDING (POTENTIAL).
FT	SITE 240 240	SUBSTRATE BINDING (POTENTIAL).
SQ	SEQUENCE 756 AA; 84561 MW; 6954F39A25E73B0A CRC64;	

Query Match	4.18;	Score 237.5;	DB 1;	Length 756;
Best Local Similarity	20.18;	Pred. No. 7.1e-09;		
Matches 156;	Conservative 89;	Mismatches 232;	Indels 301;	Gaps 31;

QY	280	VLRVLVSIFLRYLRITNVR-NAYPLMLLSVI---CEIWFSWILDQPKKFPINRETY	335
DB	80	VLSALVSLRYLWRITETFDNTFWIQIGLVILLMAELYALMFLSYFQTQPHRAP-	138
QY	336	LDRALRYDRGEESQLAAVDFIVSTVDDPKPEPITVANTVLSILAVYDPVKDVCYSVD	395
DB	139	---LPLPDNVDDWPT---VDFIPTYD---EQLSIVRLTVLGGIDWPPDKVNVYILD	188
QY	396	DGASMLTFDALAETSEFAKKWVPVKYDIETPRAPEYFCQIDYLDKVKQPSFYKDRRA	455
DB	189	DG-----VRP-----	193
QY	456	MKREYEEFKIRINALVSKALVPEEGWIMQDGTWPNGNTRDHPGMIQVFLGHSGLDTE	515
DB	194	---EPEQAFKCGAL-----YIGRVDD	213
QY	516	GNELPLRVYVSEKRPQGHHKKAGAMNALVRYSAVLTNGOYMLNDCDHYLNNSKAYRE	575
DB	214	-----SAHAKAGNLNHAIK---RTSGDYILILDCDH-IPTRAFLQI	250
QY	576	AMCFLMDPNLGPQVCYVQFORFGIDRNDRYANRNTVFEDINLRGLDGIQGFVYVGTGC	635
DB	251	AMGWVADR---KIALMQTPHHFSPD-----	274
QY	636	VFNRTAIYGYEPPIIKAKPGFLSLCGGKKKASKKRSDKKKSNKHVSDVSPVFNLED	695
DB	275	PFORNIAVGYRTP-----PGNLT--	292

Qy	696	IEEGVEGAGDEDEKSVLMQSMLEKRFQGSAAFAVSTLMEYGVGVQDSPTPSLKEAIHV	755
Ds	293	-FYGVTDGND-----FDATEFCSCAI-----LREAIRES	323
Qy	756	ISGQYEDKSWGTEIGWIYGSVTEDILTGFKMHARGWSRVYCMKPRPAKFSAPINLSDR	815
Ds	324	IG-----GFAVETVEDAHTALRMQRGWSTAYL--RIPVASGLATERLTH	368
Qy	816	LNQVLRWALGSVEILSRHCPLWYGGGRKLEFEPAYINTIYPLTSLPYLCILPAI	875
Ds	369	IGORMRWARGMIQI-FRVDPNM--LGRGLKLGKLCYLSAMTSPFFAIPRIVELASPLA	424
Qy	876	CLITGKFIEMPEISNLASIWFIALFISIFATGILEMRWS---GVGIDEWRRNEQFW-VIGG	931
Ds	425	FLFAGQNI-----AAAPLAAVAYALPHMFHSIATAAKVKNKGR-VYSEWSEVYE	472
Qy	932	ISAHLEFAVFGOLKVLVLAGIDTFTYTSKA---NDEGDEFAELY-MFKWTTLLIPTTI--	985
Ds	473	TTWALFLVIRVITVTLFPKSGKFNVTGKGVLEESEFDLGATGYPNIIFATIMGGLLIGL	532
Qy	986	--LII--NMGVVAVAGTSYAINSGYOSWGPLFGKLFFAFWVIVHLVPFLKGL-MGRQNR	1038
Ds	533	FELIVRPNQLDVIRNAYLLNCA-----WALISLIILEFAAIAVGRETK	575
RESULT 11			
BCSA_XANAC			
ID	BCSA_XANAC	STANDARD;	PRT; 729 AA.
AC	P58932;		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).		
GS	BCSA_OR XAC3516.		
GN	Xanthomonas axonopodis (pv. citri).		
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;		
OC	Xanthomonas.		
ON	NCBI_TaxID=92829;		
FX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=306 / ATCC 13902 / XV 101;		
FX	MEDLINE=22022145; PubMed=1202417;		
RA	da Silva A.C.R., Ferrto J.A., Reinach F.C., Farah C.S., Furlan L.R.,		
RA	Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,		
RA	Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,		
RA	Camaretto G., Cannavan F., Cardozo J., Chambergo F., Cipolina L.P.,		
RA	Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,		
RA	Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,		
RA	Fatighieri E.F., Franco M.C., Greggio C.C., Gruber A.,		
RA	Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,		
EA	Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,		
RA	Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,		
RA	Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,		
RA	Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,		
RA	Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,		
RA	Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,		
RA	Setubal J.C., Kitajima J.P.;		
RT	"Comparison of the genomes of two Xanthomonas pathogens with differing		
RT	host specificities."		
PL	Nature 417:459-463(2002).		
CC	-I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes		
CC	uridine 5'-diphosphate glucose to cellulose, which is produced as		
CC	an extracellular component for mechanical and chemical protection		
CC	(By similarity).		
CC	-I- CATALYTIC ACTIVITY: UDP-glucose + {(1,4-beta-D-glucosyl)}(N) = UDP		
CC	+ {(1,4-beta-D-glucosyl)}(N+1).		
CC	-I- COFACTOR: Magnesium (By similarity).		
CC	-I- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid		
CC	(c-di-GMP) (By similarity).		
CC	-I- PATHWAY: Bacterial cellulose biosynthesis.		
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane		
CC	(Potential).		
CC	-I- DOMAIN: There are two conserved domains in the globular part of		

CC the protein: the N-terminal domain (domain A) contains
CC the conserved DXD motif and is possibly involved in catalysis and
CC substrate binding. The C-terminal domain (domain B) contains the
CC QXXRW motif and is present only in processive glycosyl
CC transferases. It could be involved in the processivity function of
CC the enzyme, possibly required for holding the growing glycan chain
CC in the active site.
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE012000; AA038361.1; -
KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
KW Transmembrane; Inner membrane.
FT DOMAIN 151 244 CATALYTIC SUBDOMAIN A.
FT DOMAIN 321 381 CATALYTIC SUBDOMAIN B.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 405 425 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 520 540 POTENTIAL.
FT TRANSMEM 549 569 POTENTIAL.
FT TRANSMEM 610 630 POTENTIAL.
FT ACT_SITE 193 193 POTENTIAL.
FT ACT_SITE 337 337 POTENTIAL.
FT SITE 240 240 SUBSTRATE BINDING (POTENTIAL).
FT SITE 242 242 SUBSTRATE BINDING (POTENTIAL).
SQ SEQUENCE 729 AA; 80915 MW; B9C08BB995E795B1 CRC64;

Query Match 4.08; Score 230; DB 1; Length 729;
Best Local Similarity 19.28; Pred. No. 2.3e-08;
Matches 149; Conservative 94; Mismatches 211; Indels 324; Gaps 31;
QY 278 VIVRLVLSI-----FLHYRLTNP--VRNA--VPLWLLSVICELWFPALSWILOFPKWF 328
DB 77 VVLMGMGSLAVSCRYIWNRTQTMGVSAVDLGLGLGAELYAFVILVGFQVLM 136
QY 329 PINRTYLDRLALYDRGEPSSQAANDIFVSTVDPLKEPPIVANTVLSILAYDYPVK 388
DB 137 PLNRKP-----VPLPADQLWPS---VDVFPTYN--EPLSVVVRTVLAASVIDWPAGK 185
QY 389 VSCYVSDDGASMLTFDALAETSEFARKWVPVKYDIEPRAPEF-YFCOKIDYLDKQVP 447
DB 186 IYIHLDDG-----RDEFRAPCAEV----- 206
QY 448 SFVKDRAMKREYEEFKIRINALVSKALVPPEGWIMODGTPWPCNNTRDHPGMIOVFLG 507
DB 207 ----- 206
QY 508 HSGGLDTGNEPLRLVYVYREKRCFQHHKKAGAMNALVRYSAVLNCOYMLNDCDHYI 567
DB 207 ---GIN-----YVTRNNA---HAKAGNINAALK-----KCSGDYVAIFDCDH-I 244
QY 568 NNSKAVREAM-CFLMDPNLGPQVCYVQPPQRF---DGIDRN-----DRYANRNTVFFDINL 619
DB 245 PTRSFLQVAMGFWLHDTKL-----ALVQMPHYFFSPDPERNLDTGKVPNEGELFYGLLQ 300
QY 620 RGLDGIQGVVYVGVGVNFRPAIYGYEPPIKAKPGFLASLCGGKKKASKSKSSDKKK 679
DB 301 DGDQWNATFCGSCAVLKRTA----- 322
QY 680 SNKHVDSSVPVFNLEIDIEEGVAGAGFDEKSVLMSQMSLEKRFQSGAFAFVASTLMYEGV 739
DB 323 -----LEEYGV 329
QY 740 PQSPTPELLKEATHVISCGYEDKSEWGTETIGYGSVTEIDILYGFPMHARGWRSVCM 799

Db 330 -----AVE-----
QY 800 KRPAKFSAPINLSDRNQVLR
Db 357 --PQAAGLATESLSGHVQRI
QY 860 PLTSLPLVYCIPLPAICLLTGR
Db 411 FFYGVPRRIITAPLAYLFFGA
QY 919 EWMNEQFWIGGISAHFLFAVF
Db 471 EYVETTLAWI-----E
QY 973 FKWTLLIPPTTILIIINMGV
Db 521 F-----LULLNVGMV
RESULT 12
ACSL_ACEXY STANDARD;
ID P21877; P37717;
DT 01-MAY-1991 (Rel. 18, Creat
DT 15-JUN-2002 (Rel. 41, Last
DE Cellulose synthase 1 (Inclu
DE [UDP-forming] (EC 2.4.1.12)
DE synthase 1 regulatory domain
GN ACSAB OR ACSA OR ACSB.
OS Acetobacter xylinus.
OC Bacteria; Proteobacteria; E
OC Gluconacetobacter.
OX NCBI_TaxID=28448;
RN [1]
RP SEQUENCE FROM N.A., AND SE
RC STRAIN=ATCC 53582;
RX MEDLINE=91346705; PubMed=2
RA Saxena I.M., Lin F.C., Bro
RT "Cloning and sequencing of
gene of Acetobacter xylinu
Plant Mol. Biol. 15:673-68
[2]
RN SEQUENCE FROM N.A., AND SE
RC STRAIN=ATCC 53582;
RX MEDLINE=91322509; PubMed=1
RA Saxena I.M., Lin F.C., Bro
RT "Identification of a new ge
in Acetobacter xylinum".
Plant Mol. Biol. 16:947-95
[3]
RN REVISIONS.
RP STRAIN=ATCC 53582;
RX PubMed=8083166;
RA Saxena I.M., Kudlicka K.,
RT "Characterization of genes
operon) of Acetobacter xyl
crystallization.";
J. Bacteriol. 176:5735-575
[4]
RN SEQUENCE OF 1-8 FROM N.A.
RC STRAIN=ATCC 23769;
RX MEDLINE=94131945; PubMed=8
RA Standal R., Iversen T.-G.,
RA Valla S.;
RT "A new gene required for c
cellulolytic activity in A
bcs operon.";
J. Bacteriol. 176:665-672(
[5]
RN FUNCTION.
RP STRAIN=ATCC 53582;

RX PubMed-2138620;
 RA Lin F.C., Brown R.M. Jr., Drake R.R. Jr., Haley B.E.:
 RT "Identification of the uridine 5'-diphosphoglucose (UDP-Glc) binding
 subunit of cellulose synthase in *Acetobacter xylinum* using the
 photoaffinity probe 5-azido-UDP-Glc."; *J. Biol. Chem.* 265:4782-4784(1990).
 RL [6]
 RN 3D-STRUCTURE MODELING, AND MUTAGENESIS OF D-188; D-189; D-236; D-333;
 RP Q-369; R-370 AND R-372.
 RC STRAIN-ATCC 23769;
 RA Saxena I.M., Brown R.M. Jr., Dandekar T.:
 RT "Structure-function characterization of cellulose synthase:
 relationship to other glycosyltransferases."; *Phytochemistry* 57:1135-1148(2001).
 RL [7]
 RN REVIEW ON DOMAIN ARCHITECTURE.
 RP PubMed-7883697;
 RA Saxena I.M., Brown R.M. Jr., Fevre M., Geremia R.A., Henriessat B.:
 RT Multidomain architecture of beta-glycosyl transferases: implications
 for mechanism of action."; *J. Bacteriol.* 177:1419-1424(1995).
 RL -!- FUNCTION: Bifunctional protein comprised of a catalytic subunit
 and a regulatory subunit. The catalytic subunit of cellulose
 synthase polymerizes uridine 5'-diphosphate glucose to cellulose
 in a processive way. The thick cellulosic mats generated by this
 enzyme probably provide a specialized protective environment to
 the bacterium. The regulatory subunit binds bis-(3'-5') cyclic
 diguanylic acid (c-di-GMP).
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-beta-D-glucosyl](N) = UDP
 + [(1,4)-beta-D-glucosyl](N+1).
 CC -!- COFACTOR: Magnesium (By similarity).
 CC -!- ENZYME REGULATION: Activated by c-di-GMP.
 CC -!- PATHWAY: Bacterial cellulose biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (Potential).
 CC -!- DOMAIN: There are two conserved domains in the globular part of
 the catalytic subunit; the N-terminal domain (domain A) contains
 the conserved DXD motif and is possibly involved in catalysis and
 substrate binding. The C-terminal domain (domain B) contains the
 QXXRW motif and is present only in processive glycosyl
 transferases. It could be involved in the processivity function of
 the enzyme, possibly required for holding the growing glycan chain
 in the active site.
 CC -!- INDUCTION: Cellulose is produced at a linear rate with respect to
 cell growth when O12 is present.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 GLYCOSYLTRANSFERASE FAMILY 2.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB
 FAMILY.
 CC -!- CAUTION: Was originally (Ref.1) thought to be two separate ORFs
 named acsa and acsb, due to a frameshift in position 678.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; X54676; CAA38487.1; ALT_FRAME.
 DR EMBL; X54676; CAA38488.1; ALT_FRAME.
 DR EMBL; M96060; AAA16971.1; .
 DR PIR; S14437; S14437.
 DR PIR; S13732; S13732.
 DR InterPro; IPR003919; Cellsynth_A.
 DR InterPro; IPR001173; Glycos_transf.2.
 DR Pfam; PF00535; Glycos_transf.2; 1.
 DR PRINTS; PRO1439; CELLSYNTHASE.
 KW Cellulose biosynthesis; Transferase; Glycosyltransferase;
 Transmembrane; Inner membrane.
 FT DOMAIN 1 741 CATALYTIC.

FT	DOMAIN	742	1550	CYCLIC DI-GMP BINDING.
FT	DOMAIN	147	240	CATALYTIC SUBDOMAIN A.
FT	DOMAIN	317	377	CATALYTIC SUBDOMAIN B.
FT	TRANSMEM	26	46	POTENTIAL.
FT	TRANSMEM	47	67	POTENTIAL.
FT	TRANSMEM	106	126	POTENTIAL.
FT	TRANSMEM	398	418	POTENTIAL.
FT	TRANSMEM	423	443	POTENTIAL.
FT	TRANSMEM	468	488	POTENTIAL.
FT	TRANSMEM	507	527	POTENTIAL.
FT	TRANSMEM	547	567	POTENTIAL.
FT	TRANSMEM	1513	1533	POTENTIAL.
FT	ACT_SITE	333	333	POTENTIAL.
FT	SITE	236	236	SUBSTRATE BINDING (POTENTIAL).
FT	MUTAGEN	188	188	D->N: DECREASE IN ACTIVITY.
FT	MUTAGEN	188	188	D->P: LOSS OF ACTIVITY.
FT	MUTAGEN	189	189	D->Y: LOSS OF ACTIVITY.
FT	MUTAGEN	236	236	D->Y: LOSS OF ACTIVITY.
FT	MUTAGEN	333	333	D->R: LOSS OF ACTIVITY.
FT	MUTAGEN	369	369	Q->M: LOSS OF ACTIVITY.
FT	MUTAGEN	370	370	R->P: LOSS OF ACTIVITY.
FT	MUTAGEN	370	370	R->Q: DECREASE IN ACTIVITY.
FT	MUTAGEN	372	372	R->A: LOSS OF ACTIVITY.
SQ	SEQUENCE	1550 AA; 168161 MW; 63AB8952BC39E961 CRC64;		

Query Match 3.6%; Score 207; DB 1; Length 1550;
 Best Local Similarity 19.6%; Pred. No. 3e-06;
 Matches 154; Conservative 88; Mismatches 245; Indels 298; Gaps 29;

QY	280	VLRVLVLSIFLHYRLTNPVRNAYPLWL	-----LSVCEIWFALSWITLQDFPKWEPINRE	333
DB <td>80 <th>VLSGLVSLRYLTWRLET</th> <th>--SPDTWLGGLTMLLVAELYALMMFLSYFQFIAPLHRA</th> <th>137</th> </td>	80 <th>VLSGLVSLRYLTWRLET</th> <th>--SPDTWLGGLTMLLVAELYALMMFLSYFQFIAPLHRA</th> <th>137</th>	VLSGLVSLRYLTWRLET	--SPDTWLGGLTMLLVAELYALMMFLSYFQFIAPLHRA	137
QY <th>334</th> <th>TYLDRLALRYDREGEPSQLAAVDIFVSTVDPLKEPPIVTANTVLSILAVDPYDKVSCYV</th> <th>393</th>	334	TYLDRLALRYDREGEPSQLAAVDIFVSTVDPLKEPPIVTANTVLSILAVDPYDKVSCYV	393	
DB <th>138</th> <th>P-----LPL-----PPNPDEWPTVDIFVPTYN</th> <th>---EELSIVRLTVLGLSLGIDWPPKRVVHI</th> <th>186</th>	138	P-----LPL-----PPNPDEWPTVDIFVPTYN	---EELSIVRLTVLGLSLGIDWPPKRVVHI	186
QY <th>394</th> <th>SDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFVFCOKIDYLDKQVPSFVKDR</th> <th>453</th>	394	SDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPFVFCOKIDYLDKQVPSFVKDR	453	
DB <th>187</th> <th>LDGG-----</th> <th>-----RREF-----</th> <th>195</th>	187	LDGG-----	-----RREF-----	195
QY <th>454</th> <th>RAMKREYEEFKIRINALVSKALKVPEGWIMQDTPWGNTRDHPGMIOVFLGHSGGLD</th> <th>513</th>	454	RAMKREYEEFKIRINALVSKALKVPEGWIMQDTPWGNTRDHPGMIOVFLGHSGGLD	513	
DB <th>196</th> <th>-----</th> <th>-----AAFAECGA--</th> <th>204</th>	196	-----	-----AAFAECGA--	204
QY <th>514</th> <th>TEGNELPRLVYVSRKRPFGQHHKAGAMNALVRVSAVLNGOYMLDCHDHYINNKA</th> <th>573</th>	514	TEGNELPRLVYVSRKRPFGQHHKAGAMNALVRVSAVLNGOYMLDCHDHYINNKA	573	
DB <th>205</th> <th>-----NYIART-----NEHAKAGNLNYAIG</th> <th>---HTDGYILIFDCDH-VPTRAFL</th> <th>246</th>	205	-----NYIART-----NEHAKAGNLNYAIG	---HTDGYILIFDCDH-VPTRAFL	246
QY <th>574</th> <th>REAMCFLMDPNLGPQVCYVQFPQDFGIDRNDRYANRNTVFFDINRLGLDGIQGPVYVGT</th> <th>633</th>	574	REAMCFLMDPNLGPQVCYVQFPQDFGIDRNDRYANRNTVFFDINRLGLDGIQGPVYVGT	633	
DB <th>247</th> <th>QLTMGWVE--DPKIALMQTPHHFYSPD</th> <th>-----</th> <th>272</th>	247	QLTMGWVE--DPKIALMQTPHHFYSPD	-----	272
QY <th>634</th> <th>GCVENRTALYGVPEPIKAKKPGFLASLGGKKKASKKRSDDKKSKKHVDSSVPVFN</th> <th>693</th>	634	GCVENRTALYGVPEPIKAKKPGFLASLGGKKKASKKRSDDKKSKKHVDSSVPVFN	693	
DB <th>273</th> <th>--PFORNLSAGYRTP--</th> <th>-----PEGNL</th> <th>290</th>	273	--PFORNLSAGYRTP--	-----PEGNL	290
QY <th>694</th> <th>EDIEEGVAGDFDEKSVLMSQMSLEKRFQGSAAFAVATLMYEGGVQPSSTPESLKEAI</th> <th>753</th>	694	EDIEEGVAGDFDEKSVLMSQMSLEKRFQGSAAFAVATLMYEGGVQPSSTPESLKEAI	753	
DB <th>291</th> <th>---FYGVQDGDND-----FWDATPFCGSCAI</th> <th>-----LRRTAI</th> <th>319</th>	291	---FYGVQDGDND-----FWDATPFCGSCAI	-----LRRTAI	319
QY <th>754</th> <th>HVISCYEDKSEWGTETIGNIYGVSTEDILTGFKMHARCSRVYCMKPAKFSAPINLS</th> <th>813</th>	754	HVISCYEDKSEWGTETIGNIYGVSTEDILTGFKMHARCSRVYCMKPAKFSAPINLS	813	
DB <th>320</th> <th>EQIG-----</th> <th>---GFATQTVTEDAHTALKMORLGWSTAYL--RIPLAGGLATERLI</th> <th>364</th>	320	EQIG-----	---GFATQTVTEDAHTALKMORLGWSTAYL--RIPLAGGLATERLI	364
QY <th>814</th> <th>DRLNQVLRWALGSVEILSRHCPLMYGGYGRKLEFAYINTIYPLTSLPLVYILP</th> <th>873</th>	814	DRLNQVLRWALGSVEILSRHCPLMYGGYGRKLEFAYINTIYPLTSLPLVYILP	873	
DB <th>365</th> <th>LHIGQRVWARGLQI-FRIDNPL---</th> <th>---FGRGLSGORLCYLSAMTSFLEAVPRVIFLSSP</th> <th>420</th>	365	LHIGQRVWARGLQI-FRIDNPL---	---FGRGLSGORLCYLSAMTSFLEAVPRVIFLSSP	420
QY <th>874</th> <th>ACLTLTGKFIEMPEISNLSIASIFALFISFATGILEMRWSCVG-----IDENWRNEQW-V</th> <th>928</th>	874	ACLTLTGKFIEMPEISNLSIASIFALFISFATGILEMRWSCVG-----IDENWRNEQW-V	928	


```

Db 935 AAVEPANTNDKGEYMLEAVEGA-----YTIKVAAPGY----- 967
QY 452 DRAMKREYEFEKIRINALVSK--ALK--VPEEGTMDQ-----GTPWPGNN 494
Db 968 -----SDEFSVELKGVDTKRETKALPKVPVGPGETAYDDGTAENANSYFAAGNGWAVKM 1019
QY 495 T-----RDHPGMIQVFLGHSGGL-----DTE-----GNELPRLY--VSREKRPFGQHKKKA 539
Db 1020 TLADGKD-KGML-----TGGLFRFWDTEFPDPGGTEFKVEYVDATGKDGAPG---KKIA 1069
QY 540 GAMNALVRVSAVLNCGYM-----LNLDCDHYINNSKAVREAMCFMLDNPGLPQVCY 591
Db 1070 GPENA-----EALRNGEWTKVDLSSKGINVDKDFYL-----VY 1102
QY 592 VQP-PORFD---GIDRNDRYANRNTVFEDINLRGLDGIQGVVVGTCVFNRTAIVGYE- 646
Db 1103 IQSKPDYPSPGLAMDETQNSGRNWQYIDGKWQPGDKADGNYMI-----RALVDYEA 1154
QY 647 --PPIKA-----KKPGFLASLCGGKKKASKSK----- 671
Db 1155 AVPEITSPTDKSYTNKDSVTVKGNASPGTTHVYNGEKEAGETKAAADGTFHAGILNKG 1214
QY 672 --KRSSDKKKSKNHVDSSVPV-----ENLEDIEEG-----VEGAGFDDE-K 709
Db 1215 ENELTATADNGTDTASSPIVTLDQEKPELTLDNPKDGGKTKNETLTIVKGAVSDNLK 1274
QY 710 SVLMSOMSLKRFQGSAAVASTIMEYG 737
Db 1275 DV---KVNGKKATVADGYSARILLENG 1299

RESULT 15
YCDQ_ECOLI
ID YCDQ_ECOLI STANDARD; PRT; 441 AA.
AC P75905;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ycdQ.
GN YCDQ OR B1022.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Bacterichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -|- SIMILARITY: TO Y.PESTIS HMS LOCUS PROTEIN HMSR AND TO
CC S.EPIDERMIDIS IC9A.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```

CC between the Swiss Institut
CC the European Bioinformati
CC use by non-profit institut
CC modified and this stateme
CC entities requires a licens
CC or send an email to licens
CC -----
DR EMBL; AE000204; AAC74107.1;
DR EMBL; D90739; BAA35803.1;
DR EcoGene; EGI3863; ycdQ;
DR InterPro; IPR001173; Glycos
Pfam; PF00535; Glycos.Trans
KW Hypothetical protein; Trans
FT TRANSMEM 5 25
FT TRANSMEM 33 53
FT TRANSMEM 331 351
FT TRANSMEM 363 383
FT TRANSMEM 395 415
FT SEQUENCE 441 AA; 50765
SQ
Query Match 2.08
Best Local Similarity 18.1%
Matches 92; Conservative
QY 482 WIMOD-GTPW----PGNTRD
Db 53 WYERHWPWGENAPAPQLKD
QY 510 GG-----LDTEGNELP--R
Db 113 GSTDKTRAILDRMAAQIPLR
QY 561 LDCDHYINNSKAVREAMCFLM
Db 162 IDGDALLD-----RDAAAYIV
QY 621 GLDGIQGVVYG-----TGCV
Db 201 ----LVGKIQVGEYSIIIGLI
QY 676 DKKSKNHVDSSVPVFNLEDI
Db 226 -----NVFTVS--
QY 736 YGVFQPSSTPESLLKEAHVI
Db 245 VG-----
QY 796 YCMPKRPAPKGSAPINLSDR
Db 271 FVEPRALCWI-LMPETLKG
QY 856 TTIYPLTSL-PLLYVCI-----
Db 329 TTIWAFTCLVGFIIYAVQLAG
QY 888 --SNLAS-----IWFIALF--
Db 387 YEHNLTSSLFWIWFPPVFW
Search completed: February 19,
Job time : 59 secs
```

```

Bioinformatics and the EMBL Outstation -
stitute. There are no restrictions on its
as long as its content is in no way
ot removed. Usage by and for commercial
ement (see http://www.isb-sib.ch/announce/
sib.ch).
-----
sf_2.
1.
ane; Complete proteome.
ENTIAL.
ENTIAL.
ENTIAL.
ENTIAL.
ENTIAL.
34F5A53D72FEBABB CRC64;
ore 116.5; DB 1; Length 441;
sd. No. 1.2;
Mismatches 130; Indels 229; Gaps 23;
-----
IIIPCNEEKVVEETIIHAALAOYENIEVIAVD 112
REKRPGFQHHKKAGAMNALVRVSAVLNCGYMLN 560
-----NQGKAIAL-----KTGAAAKSEYLVLC 161
POVCYVQPPQFDGIDRNDRYANRNTVFEDINLR 620
-----YNPRVGAVTGNPRTRST----- 200
UYGYEPPKAKKPGFLASLCGGKKKASKSKRSS 675
IYG-----
GAGFDDEKSVLSMSQMSLEKRFQCSAAVASTIME 735
-----GVIAAFRRSALAE 244
DKSEWGTETIGWYGSVTEIDLTGPKMHARGHRSV 795
-----YNSDDMITEDIDISWKLQLNQM-TI 270
SWALGSVEILFSRHCPWYGYGRLKFLERFAYIN 855
SWAOGGAEEVFLKNMTRLWRKENFRMWPL-FFEYCL 328
-----LPAICLLTGKFMPEI----- 887
CELTHIAATHTAGILLCTICLL--OPIVSLMIENR 386
AT 905
TT 413
```

Result No.	Query			ID	Description
	Score	Match	Length		
1	5423.5	93.9	1079	10	Q9L1T1 zea mays (m
2	5421.5	93.8	1077	10	Q9L1T6 zea mays (m
3	5143	89.0	1067	10	Q9L1T5 zea mays (m
4	4577.5	79.2	1067	10	Q9XG36 gossypium h
5	4530	78.4	1065	10	Q9FHK6 arabidopsis
6	4507	78.0	1065	10	Q9FHK6 arabidopsis
7	3993.5	69.1	1074	10	Q9L1T8 zea mays (m
8	3981	68.9	1075	10	Q9L1T9 zea mays (m
9	3963.5	68.6	1081	10	Q48946 arabidopsis
10	3844.5	66.5	1094	10	Q9L1T2 zea mays (m
11	3820	66.1	1059	10	Q9L1T4 zea mays (m
12	3818	66.1	1086	10	Q9L1T3 zea mays (m
13	3801	65.8	1026	10	Q9SWM6 arabidopsis
14	3795	65.7	1026	10	Q9XHP6 arabidopsis
15	3778.5	65.4	1065	10	Q9SKJ5 arabidopsis
16	3746	64.8	1043	10	Q9FNC3 arabidopsis

Db	61	CPOCKNKKYKRRHKGSPAIRGEEDDTDADSDAFNYPASGNDQKQKIADMRSRWRNAGG	120
Qy	121	SGNVGHPKYDSGEIGLSKYSDEIRGYPVSTNSOMSGEIGASPDHMHMSPTGNISR	180
Db	121	SGDVGPRKYDSGEIGLTKYDSGEIRGYPVSTNSQISGEIGASPDHMHMSPTGNIGRR	180
Qy	181	APFPYVNHSPNSRFSFGSGIGNVANKERVDGWMKQDKGAIPTWNTGTSIAPSEGRAATI	240
Db	181	APFPYVNHSSNSRFSFGSGVGNVANKERVDGWMKQDKCTIPMTNGTSIAPSEGRGVGI	240
Qy	241	DASTEYNMEDALLNDETROPLSRKVPDIASKINPYRMVTVLRVLVLSIFLHYRLTNPVNR	300
Db	241	DASTDYNMEDALLNDETROPLSRKVPDIASKINPYRMVTVLRVLVLSIFLHYRLTNPVNR	300
Qy	301	AYPLWLLSVICETWFAWLSWLDQFPKWFPIINRETYLDRLALRYDREGESQLAADVFS	360
Db	301	AYPLWLLSVICETWFAWLSWLDQFPKWFPIINRETYLDRLALRYDREGESQLAADVFS	360
Qy	361	TVDPLKEPPIVANTVLSILAVDYPVDKYSVSDGASMLTFDALAETSEFARKWVPFV	420
Db	361	TVDPMKEPPLVANTVLSILAVDYPVDKYSVSDGAAMLTFDALAETSEFARKWVPFV	420
Qy	421	KYDIEPRAPFVFCQKIDYLDKQVOPSVFKDRAMKREYEFEKIRINALVSKALKVPEE	480
Db	421	KYINIEPRAPFVFSQKIDYLDKQVOPSVFKDRAMKREYEFEKIRVNGLVAKAQVPEE	480
Qy	481	GMTODGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSRKRPFGFHHKAG	540
Db	481	GMTODGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSRKRPFGFHHKAG	540
Qy	541	AMNALVRVSAVLINGOYMLNDCDHYINNSKAVREAMCFMLDPNLPQVYVQFPQRFDG	600
Db	541	AMNALVRVSAVLINGOYMLNDCDHYINNSKALREAMCFMLDPNLPGRVCYVQFPQRFDG	600
Qy	601	IDRNDRYANRNTVFFDINLRGLDIOGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL	660
Db	601	IDRNDRYANRNTVFFDINLRGLDIOGPVYVGTGCVFNRTALYGEPPKQKKGFLSSL	660
Qy	661	CGGKKKASKSKRSDDKSNKHVDSSVPFNLEDIEEGVEGAGFDDKSVLMSQMSLEK	720
Db	661	CGGRKKGSKSKK-GSDKKSQKHVDSSVPFNLEDIEEGVEGAGFDDKSLMSQMSLEK	719
Qy	721	RFQSAAFVASTLMEYGGVQPSSTPSLLKEAITHVISCYEDKSEWGTEIGWYGSVTE	780
Db	720	RFQSAAFVASTLMEYGGVQPSATPSLLKEAITHVISCYEDKIEWGTEIGWYGSVTE	779
Qy	781	ILTGFKMHARGWSVCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLWYG	840
Db	780	ILTGFKMHARGWSIYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLWYG	839
Qy	841	YGRKLFLERFAYINTIYPLTSLPLLYCILLPAICLLTGKFTMPEISNLASLWFTALFL	900
Db	840	YGRKLFLERFAYINTIYPLTSLPLLYCILLPAICLLTGKFTIPEISNPFASLWFLSLFI	899
Qy	901	SIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA	960
Db	900	SIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA	959
Qy	961	NDEEGDFAELMYFKWTTLLIPPTILLIINNMGVYVAGTSYAINSGYQSWGLPFGKLPFAFW	1020
Db	960	SDEGDFAELMYFKWTTLLIPPTILLIINNMGVYVAGTSYAINSGYQSWGLPFGKLPFAFW	1019
Qy	1021	VIVHLYPFLKGLMGONRPTIIVVWAILLASIFSLWLWVDPFTTTLRAGPNQTCGINC	1080
Db	1020	VIVHLPPFLKGLMGONRPTIIVVWAILLASIFSLWLWVDPFTTNRVTGPDTRTCGINC	1079
RESULT	2		
ID	Q9LLI6	PRELIMINARY; PRT; 1077 AA.	
AC	Q9LLI6;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		

DT	01-JUN-2002 (TrEMBLrel. 21, annotation update)		
DE	Cellulose synthase-4.		
GN	CESA-4.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; S		
OC	Spermatophyta; Magnoliophyt		
OC	Panicoidae; Andropogoneae;		
OX	NCBI_TaxID=4577;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20398328; PubMed=10		
RA	Holland N., Holland D., Hel		
RA	Xoconostle-Cazares B., Delm		
RT	"A comparative analysis of		
RT	family.";		
RL	Plant Physiol. 123:1313-132		
DR	EMBL; AF200528; RAF89964.1;		
DR	InterPro; IPR005150; Cellul		
DR	InterPro; IPR001841; Znf_r1		
DR	Pfam; PF03552; Cellulose_sy		
DR	SMART; SM00184; RING; 1.		
SQ	SEQUENCE 1077 AA; 120501		
	5F960E4AA753E2D6 CRC64;		
	Query Match 93.8%;		
	Best Local Similarity 92.4%;		
	Matches 998; Conservative		
QY	1	MDGDADALSKRGAGDVQCIC	60
DB	1	MEDGADGVKSGRGGGVQCIC	60
QY	61	CLQCKTKYKRHRGSPAIRGEEG	120
DB	61	CPOCKTKYKRHKGSPAIRGEEG	118
QY	121	SGNVGHPKYDSGEIGLSKYDSG	180
DB	119	SGDVGPRKYDSGEIGLTKYDSG	178
QY	181	APFPYVNHSPNSRFSFGSGIGN	240
DB	179	APFPYVNHSPNSRFSFGSGIGN	238
QY	241	DASTEYNMEDALLNDETROPLS	300
DB	239	DASTDYNMEDALLNDETROPLS	298
QY	301	AYPLWLLSVICETWFAWLSWILD	360
DB	299	AYPLWLLSVICETWFAWLSWILD	358
QY	361	TVDPLKEPPIVANTVNTVLSILAV	420
DB	359	TVDPMKEPPLVANTVNTVLSILAV	418
QY	421	KYDIEPRAPFVFCQKIDYLR	480
DB	419	KYINIEPRAPFVFSQKIDYLR	478
QY	481	GMTODGTWPNGNTRDHPGMI	540
DB	479	GMTODGTWPNGNTRDHPGMI	538
QY	541	AMNALVRVSAVLINGOYMLNLD	600
DB	539	AMNALVRVSAVLINGOYMLNLD	598
QY	601	IDRNDRYANRNTVFFDINLRGI	660
DB	599	IDRNDRYANRNTVFFDINLRGI	658
QY	661	CGGKKKASKSKRSDDKSKNK	720

```
Db 659 CGGRKKASKSK-KGSDKKSKQKHDVSSVPVFNLEIDIEBGEVBCAGFDDEKSLMSQMSLEK 717
QY 721 RFQSAFVASTLMVEYGVQSSPTESILLKAIHVISGVEDKSEWTEIGWIYGSVTE 780
Db 718 RFQSAFVASTLMVEYGVQSSPTESILLKAIHVISGVEDKSEWTEIGWIYGSVTE 777
QY 781 ILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRLNOVLWALGSVEILFSRHCPLWYG 840
Db 778 ILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRLNOVLWALGSVEILFSRHCPLWYG 837
QY 841 YGRLKFLERFAVINTTIYPLTSLPLLYCYILPAICLTGKFIPEISNLASIWFIALFL 900
Db 838 YGRLKFLERFAVINTTIYPLTSLPLLYCYILPAICLTGKFIPEISNLASIWFIALFL 897
QY 901 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHFLFAVFGGLKVLGADIDTNTFTVSKA 960
Db 898 SIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHFLFAVFGGLKVLGADIDTNTFTVSKA 957
QY 961 NDEEGDFAELYMFKWTLLIPPTTILINMGVWVAGTSYAINSGYQSGWGLFGKLFPAFW 1020
Db 958 SDEGDFAELYMFKWTLLIPPTTILINMGVWVAGTSYAINSGYQSGWGLFGKLFPAFW 1017
QY 1021 VIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRLAGENIQTGINC 1080
Db 1018 VIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRLAGENIQTGINC 1077

RESULT 3
ID Q9LLI5 PRELIMINARY; PRT; 1076 AA.
AC Q9LLI5;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxId=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20398320; PubMed-10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoonostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (CesA) gene
family."
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200529; AAF89965.1;
DR InterPro; IPR005150; Cellulose_synt.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1076 AA; 120724 MW; 41E7C20EDA54F27 CRC64;

Query Match 89.0%; Score 5143; DB 10; Length 1076;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 947; Conservative 69; Mismatches 59; Indels 6; Gaps 6;

QY 1 MDGDADALSGRHGAGDVCOICADGLGTLDGDVFTACDVCRFPVPCYCEHERKEGTQA 60
Db 1 MDG-GDATNSGRHAGVQCQICGDGVTGAADGDLFTACDVGCFPCYCEYERKDGTOA 59
QY 61 CLQCKTKYKRRHRSIPAIREGDDTDADDGSDFNYPASGTEDQOKIADMRWSRMTGG 120
Db 60 CPCKTKYKRRHRSPPVHGEENEDVDADDVSDNYQASGNODQOKIAERMLTWTNRSG 119
QY 121 SGNVGHPKYDSGEIGLSKYDSGEIPRGVPSVTNSQMSGEITPGASPDHMMSPGNTSRR 180
Db 120 S-DIGLAKYDSGEIGHGKYDSGEIPRGVIPSLSQISGEITPGASPD-HMMSVPGNIGRR 177
QY 181 A-PFPVYVNHSPNSPREFSSIGNVAWKERVGDGWMKQKGAIPWTNGTSTIAPSEGRATD 239
```

```
Db 178 GHQFPVYVNHSPNSPREFSSIGNVAWKERVGDGWMK-DKGAIPWTNGTSTIAPSEGRGVA 236
QY 240 IDASTYENMEDALLNDETROPLSRKVPYPTASSKINPYRMVIVLRLVVLISLHRLTNPVR 299
Db 237 IDASTYENMEDALLNDETROPLSRKVPYPTASSKINPYRMVIVLRLVVLISLHRLTNPVR 296
QY 300 NAYPLWLLSVICEIWFALSWILDQFPKWFPPINRETYIDRLALRYDREGEPSQLAAVDIFV 359
Db 297 NAYPLWLLSVICEIWFALSWILDQFPKWFPPINRETYIDRLALRYDREGEPSQLAPVDIFV 356
QY 360 STVDPKLEPPIVTANTVLSILAVDPYDKVSCYVSDDGASMLTDLALAESEFARKWVPF 419
Db 357 STVDPKLEPPIVTANTVLSILAVDPYDKVSCYVSDDGASMLTDLALAESEFARKWVPF 416
QY 420 VKYDIEPRAPFEYFCQIDYLDKQVQSPFVKDRRAKREYEEFKIRINALVSKALKVPE 479
Db 417 CKKYNIEPAPDWPYFAQIDYLDKQVQSPFVKDRRAKREYEEFKIRINALVSKALKVPE 476
QY 480 EGWIMQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPRLVYVSRKRPQFQHHKA 539
Db 477 EGWIMQDGTTPWGNTRDHPGMIQVFLGHSGLDTEGNEPRLVYVSRKRPQFQHHKA 536
QY 540 GAMNALVRYSAVLTNGQYMLNLDGDHYINNSKAVREACFLMDPNLGPQVYVQFPORF 599
Db 537 GAMNALVRYSAVLTNGQYMLNLDGDHYINNSKALREACFLMDPNLGNVYVQFPORF 596
QY 600 GIDRNDRYANRNTVFFDINLRLGDIQGPVYVTCGVFNRTAIYGYPPPKAKKPGFLAS 659
Db 597 GIDRNDRYANRNTVFFDINLRLGDIQGPVYVTCGVFNRTAIYGYPPPKAKKPGFSS 656
QY 660 LCGGKKASKSKSKSKKSKKHVDSVPVFNLEIDIEBGEVBCAGFDDEKSLMSQMSLE 719
Db 657 LCGGKKASKSKSKSKKSKKHVDSVPVFNLEIDIEBGEVBCAGFDDEKSLMSQMSLE 715
QY 720 KRFQSAFVASTLMVEYGVQSSPTESILLKAIHVISGVEDKSEWTEIGWIYGSVTE 779
Db 716 KRFQSAFVASTLMVEYGVQSSPTESILLKAIHVISGVEDKSEWTEIGWIYGSVTE 775
QY 780 DILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRLNOVLWALGSVEILFSRHCPLWY 839
Db 776 DILTGFKMHARGWRSVYCMKRPAPKGSAPINLSDRLNOVLWALGSVEILFSRHCPLWY 835
QY 840 YGGRKFLERFAVINTTIYPLTSLPLLYCYILPAICLTGKFIPEISNLASIWFIALFL 899
Db 836 YGGRKFLERFAVINTTIYPLTSLPLLYCYILPAICLTGKFIPEISNLASIWFIALFL 895
QY 900 LSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHFLFAVFGGLKVLGADIDTNTFTVSK 959
Db 896 LSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHFLFAVFGGLKVLGADIDTNTFTVSK 955
QY 960 ANDEEGDFAELYMFKWTLLIPPTTILINMGVWVAGTSYAINSGYQSGWGLFGKLFPAF 1019
Db 956 ANDEEGDFAELYMFKWTLLIPPTTILINMGVWVAGTSYAINSGYQSGWGLFGKLFPAF 1015
QY 1020 WVIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRLAGENIQTGGIN 1079
Db 1016 WVIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRLAGENIQTGGIN 1075
QY 1080 C 1080
Db 1076 C 1076

RESULT 4
ID Q9XGX6 PRELIMINARY; PRT; 1067 AA.
AC Q9XGX6;
OC 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN CELA3.
```

OS	Gossypium hirsutum (Upland cotton).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX	eurosid II; Malvales; Malvaceae; Gossypium.
NCBI_TaxID=3635;	
[1]	
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. TEXAS MARKER-1; TISSUE=COTTON FIBER;
RX	MEDLINE=J098737; PubMed=10631273;
RA	Laosinchai W., Cui X., Brown R.M. Jr.;
RT	"A full length cDNA of cotton cellulose synthase has high homology
RT	with the Arabidopsis RSW1 gene and the cotton Ce1al gene (Accession
RL	No. AF200453) (PCR 00-002).";
RL	Plant Physiol. 122:291-291(2000).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. TEXAS MARKER-1; TISSUE=COTTON FIBER;
RA	Kimura S., Laosinchai W., Itoh T., Cui X., Brown R.M. Jr.;
RT	"Immunogold Labeling of Rosette Terminal Cellulose Synthesizing
RT	Complexes in a Vascular Plant (Vigna angularis).";
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF150630; AAD39534.2; -.
DR	InterPro; IPR005150; Cellulose_synt.
DR	InterPro; IPR001841; Znf_ring.
DR	Pfam; PF03552; Cellulose_synt; 1.
DR	SMART; SM00184; RING; 1.
SQ	SEQUENCE 1067 AA; 119325 MW; 0BBA2ED0059F29C CRC64;
	Query Match 79.2%; Score 4577.5; DB 10; Length 1067;
	Best Local Similarity 78.6%; Pred. No. 0;
	Matches 853; Conservative 100; Mismatches 105; Indels 27; Gaps
QY	2 DGDADALSGRHGAGDVQCICADGLGTITLDGVFTACDVCRPVCPCYEHKEKGTQAC 61
DB	: : : : : : : : : : : : : : :
QY	4 EGDGG-KPMKNLGQCCTCGSNVGKNTDGDPFTACNICAFPVCPCYEYRKDNQSC 62
QY	62 LCKTKYKRHRGSIPAIRGEEDGDTADDG-SDFNTPASGTEDDQKQTADMRSWRMNTGG 120
DB	: : : : : : : : : : : : : :
QY	63 PCKTRYKWQKGPSAILLGRDRETGDADDCASDFIY--SENGEQOKLAERMGNNAKYGR 120
QY	121 SONVGHPKYDSGEIGLSKYDSGELIPRGYPVSVTNSQ-MSGEIPGASPDHHMMSPGTNISR 179
DB	: : : : : : : : : : : : : :
QY	121 GDVGAPTYDK-----EISNHPLITSGQEVSGELSAASPERUSMASPGVAGG 169
QY	180 RAPFFPVYNHSPNPSREF--SGSIGNVAWKERVGDGMKKDKGAIPMTNGTSTAPSEGRAAT 238
DB	:: : : : : : : : : : : : : :
QY	170 KSSIRVV----DPVREFGSSGLGNVAWKEVRVDGMKKDKETVPMS--TCQATSE-RGLG 222
QY	239 DIDASTEYNMEDALLNETROPILSRKVPITASCKINPYRWIVTLRVLVSLFLHYRTNPV 298
DB	: : : : : : : : : : : : : :
QY	223 DIDASTDVLVDSDQLNDEARQPILSRKVSSSKINPYRWIVTLRVLVLCIFLYHRTNPV 282
QY	299 RNAYPLLLSVICIEWFALSILDQPKFWPFINRETYLDRLALRYDREGEPQSAAVDIF 358
DB	: : : : : : : : : : : : : :
QY	283 PNAYALWLISVICIEWFALSILDQPKFWPNRNITYLDRLALRYDREGEPSELAAVDIF 342
QY	359 VSTDVDPLEPPIVTANVLSILAADVDPVKVSCYSDDGASMLTFDALATSEFAKKWP 418
DB	: : : : : : : : : : : : : :
QY	343 VSTDVDPLEPPIVTANVLSILAADVDPVKVSCYSDDGAMLFLEALSETSEFAKKWP 402
QY	419 FVKYDIPEPAPEFYCQKIDYLKQKVQSFVKDRAMKREYEFBKTRINALYSKALKVP 478
DB	: : : : : : : : : : : : : :
QY	403 FCCKYNIPEPAPEWYFAQIDYLKQKVQTSFFYKDRAMKREYEEFKYRINGLVAKAQKVP 462
QY	479 ESGWTMDGTWPWCNNTRDHPGMIQVFLEGSHGLDTEGNELPLRVVYSREKRPGFQHKK 538
DB	: : : : : : : : : : : : : :
QY	463 ESGWTMDGTWPWCNNTRDHPGMIQVFLQSGGLDAEGNELPLRVVYSREKRPGFQHKK 522
QY	539 AGAMNALVRYSAVLTNGCYMLNDCDHYTNNSKAVREAMCFIMDPNLGPQVCYVQFPQR 598
DB	: : : : : : : : : : : : : :
QY	523 AGAMNALVRYSAVLTNGAFLLNDCDHYTNNSKALUREAMCFIMDPNLGKQVCYVQFPQR 582
QY	599 DGIIDNRDYANRRNVTFEDINLRGLDGIQGPVVYGTGCVFNRNTAIYGYEPTIKAK--KPQG 656

[illegible]


```
Db 599 DINRGLDGIQGVYVGTGCVFNRTALYGEPPIKVHKKPSLLSKLGGSRKKNKAKK 658
Qy 674 SSDKKSNKHVDSVPVFNLEDEEGVEGAGFDDKSVLMSQMSLEKRRFGQSAFAVASTL 733
Db 659 ESDKKSGRHTDSTVPVFNLDDEEGVEGAGFDDKALLMSQMSLEKRRFGQSAFAVASTL 718
Qy 734 MEYGGVQSSTPSELKKAHIVISCGYEDKSEWGTETGWIYGSVTEIDILTFKMHARGWR 793
Db 719 MENGVPSPATPENLLKKAHIVISCGYEDKSDWMEIGWIYGSVTEIDILTFKMHARGWR 778
Qy 794 SVYCMKRPAPKSGAPINLSDRLNOVLWALGSEIIFSRHCPULWYGGRLKFLERFAY 853
Db 779 SIYCMPLKPAKSGAPINLSDRLNOVLWALGSEIIFSRHCPULWYGGRLKFLERFAY 838
Qy 854 INTTYPLTSLPLVYICILPAICLLTGKFIEMPEISNLASIFWIALFSLFATGILEMRWS 913
Db 839 VNTTYPTISPLMLYCTLLAVCLFTNQFIPIQISNIAISWFLSLFSLFATGILEMRWS 898
Qy 914 GVGDEWRNEQFVWIGVSAHLFAVFGILLKVLGIDTNTVTYSKANDDEGDAELMYF 973
Db 899 GVGDEWRNEQFVWIGVSAHLFAVFGILLKVLGIDTNTVTYSKANDDEGDAELMYF 958
Qy 974 KWTLLIPPTLLIINMVGWVAGTSYAINSGYQSWGFLGKLFPAFWIVHLIYFPLKGLM 1033
Db 959 KWTLLIPPTLLIINMVGWVAGTSYAINSGYQSWGFLGKLFPAFWIVHLIYFPLKGLM 1018
Qy 1034 GRQNRPTIIVWAVLASIFSLWVRVDPFTTBLAGPNTOTCGINC 1080
Db 1019 GRQNRPTIIVWVSLASIFSLWVRIDPFTSVTGPDILOCGINC 1065

RESULT 7
Q9LLI8 PRELIMINARY; PRT; 1074 AA.
AC Q9LLI8;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Cellulose synthase-2.
GN CESA-2.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland B., Helentjaris T., Dhugga K.S.,
RA Xooconostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (Cesa) gene
family.";
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL: AF200526; AAF89962.1; -.
DR InterPro: IPR005150; Cellulose_synt.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synt; 1.
DR SMART: SM00184; RING; 1.
SQ SEQUENCE 1074 AA; 121187 MW; A116E66F0564E210 CRC64;

Query Match 69.1%; Score 3993.5; DB 10; Length 1074;
Best Local Similarity 68.4%; Pred. No. 1.1e-314;
Matches 749; Conservative 134; Mismatches 153; Indels 59; Gaps 20;

Qy 2 DGDADA-LKSGRHGAGDVQCICADGLGTTLDGVDFTACDVCRRFPVPCYERKEGTQA 60
Db 23 DGDAPVPAKPTKSGANGVQCICGDTGVSATGDGVFACNECAFPVPCYERKEGNQC 82
Qy 61 CLQCKTKYKRRGSPALREGECDTADD-GSDENYASGTEDQKQIADRMRSWRMTG 119
Db 83 CPQCKTRYKRGKSPRVHGDD-EEDVDLDLNEFN-----KQ----- 119
Qy 120 GSGNVGHPK----DSGEIGLS-KYDSGEIPRGVPSVTN-SQMSGEIPASPDH-MMS 172
```

```
Db 120 --GNGKPEWLOQDDADLSSS
Qy 173 PTGN-ISRRAFPVFNHSPNS
Db 174 PTSSYVDPSPVPV--RIVDP
Qy 231 PSEGRATDIDASTEYNKDAL
Db 227 -KYPEARGDME-GTGSNGEDMC
Qy 291 HYRLTNVRNAYPLWLLSVIC
Db 285 QYRISHPVRNAYGLWLVSVIC
Qy 351 QLAADVIFSVTDPLKEPPIVT
Db 345 QLAPIDVFVSTVDPLKEPLIT
Qy 411 EFARKWTFVFKYKKDIEPRAPE
Db 405 EFARKWTFVFKYKKDIEPRAPE
Qy 471 VSKALKVPEBGWIMQDCTWPG
Db 465 VAKAQVPEBGWIMQDCTWPG
Qy 531 PGFOHKKKAGAMNALRVSAV
Db 525 PGFOHKKKAGAMNALRVSAV
Qy 591 YVQFPQRFQDIDRNDRYANRT
Db 585 YVQFPQRFQDIDRNDRYANRT
Qy 651 AK--KPGFLASLGGCKKASKS
Db 645 EADLEPNIVVKSCGRRK-RKN
Qy 709 KSVLMSOMLEKREFGQSAFVA
Db 702 RSVLMSRKLEKREFGQSPFIA
Qy 769 EIGWIYGSVTEIDILTFKMHAR
Db 762 EIGWIYGSVTEIDILTFKMHAR
Qy 829 ILFSRHCPULWYGGRLKFLER
Db 822 ILLSRHCPULWYGGRLKFLER
Qy 889 NLASIFWIALFSLFATGILER
Db 882 NYAGMEFILLFASIFATGILER
Qy 949 GIDTNFTVTSKANDDEGDAE
Db 942 GIDTNFTVTSKANDDEGDAE
Qy 1009 GPLFGKLFPAFWIVHLIYFPL
Db 1002 GPLFGKLFPAFWIVHLIYFPL
Qy 1066 TRLAGPNIOTCGINC 1080
Db 1062 QKAAA--LGQCGVNC 1074

RESULT 8
Q9LLI9 PRELIMINARY; PRT; 1075 AA.
AC Q9LLI9;
DT 01-OCT-2000 (TReMBLrel. 15
ated)
```



```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase-1.
GN CESA-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoonostlie-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (CesA) gene
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200525; AAF89961.1; -.
DR InterPro; IPR005150; Cellulose_synt.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR SEQUENCE 1075 AA; 121181 MW; 67A4FBD97A811F33 CRC64;
Query Match 68.9%; Score 3981; DA 10; Length 1075;
Best Local Similarity 68.5%; Pred. No. 1.1e-313;
Matches 747; Conservative 132; Mismatches 163; Indels 48; Gaps 19;
QY 2 DGDA-DALKSGRHGAGDVOCICADGLGTTLDGDFVFTACDVCRFPVPCPYEHERKEGTOA 60
DB III : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
QY 23 DGDVPGSAKPTKSANGVCQICGSGVSATGDVFVACNECAFPCVPCYERKEGNGC 82
QY 61 CLOCKTKYKRGSPATRGEGDDTDADD-GSDFNPASGTEDQKQKIADRMRSWRMTNG 119
DB IIII : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 83 CPQCKTRYKKQKSPRVHGDE-DEEDVDDLDNEFNY-----KQ----- 119
QY 120 GSGNVGHPKYDSIGLSKYSDEIPRGYVPSVTN-SQMSGEIPGASPDHH-MMSPTGN- 176
DB III : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 120 GSGKGPWQLOGDADLSSARHE-PHRIPLRTSGQISGEIPDASPDHRSIRTSY 178
QY 177 ISRRAPPYVNHSPNSREF-SSSIGNVAWKRVDGWMKKQDKGATPMNGTSPASEGR 235
DB : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 VDPSPVPV--RIVDPSKDLNSYGLASVDWKERVESWRVKQDKNMQVFN---KYPEAR 232
QY 236 AATDIDASTEYNEDEALLDETPOLSRKVPIASSKINPVYVILVRLVLSIFLHYRLT 295
DB : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 GG-DME-GTGSNGEXMQVDARPLSLRIVPISNQLNLYRVVILRLILCFEYFVYS 290
QY 296 NPVRNAYPLWLLSVICBFWALSWLDQFPKFPINRETYLDRALRYDREGEPSQAAV 355
DB IIII : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 291 HPVRDAYGLWLVSVICBFWALSWLLDQFPKFPINRETYLDRALRYDREGEPSQAPI 350
QY 356 DIFVSTVDPLKEPPIVNTANVLSLAVDYVDKVCVSDGASMLTFDALAETSEFARK 415
DB : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 DVFVSTVDPLKEPPIVNTANVLSLAVDYVDKVCVSDGASMLTFESLSTAFARK 410
QY 416 WVPVKYDIETRAPPEYFCQIDYLDKQVQSFVKDRAMKREYEEFKIRINALYSKAL 475
DB IIII : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 411 WVPCKKHNIETRAPPEYFAQIDYLDKQIQSPFVKERRAMKREYEEFKVRINALYAKAQ 470
QY 476 KVPBEGIMQDGTWPQNNTRDHPGMIQVFLGSHSGGLDTBGNLPLRVYVSRKRPGFQH 535
DB IIII : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 471 KVPBEGTMDAGTAWPQNNTRDHPGMIQVFLGSHSGGLDTBGNLPLRVYVSRKRPGFQH 530
QY 536 HKKAGAMNALRVYSVLTNGOYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYVQPP 595
DB IIII : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 531 HKKAGAMNALRVYSVLTNGAYLLNVDCDHYFNSSKALREAMCFMMDPALGRKTCVQPP 590
QY 596 QRFQGDIDRNDRYANRNTVFEDINLRGLDGIQGPVYVGTGCVFNRTAIYGEYPIKAK--K 653
DB IIII : : : : : IIII : : : : : IIII : : : : : IIII : : : : :
DB 591 QRFQGDIDRNDRYANRNTVFEDINMKGLDGIQGPVYVGTGCVFNRTAIYGEYPIKAK--K 650
QY 654 PGFLASLCGGKKKASKSKSSDKKSNKHVDSSVPVFNLEDIEEGVEGAGDFDEKSVLM 713
```

[illegible]

QY 348 EPSOLAADVIFVSTVDLPKEPIVNTANTVLSILAVDYPVDKVCYVSDGASMLTFDALA 407
 Db 353 QFSQLAPIDFFVSTVDLPKEPIVNTANTVLSILSVDPYVERKVCYVSDGAAMLTFBAL 412
 QY 408 ETSEFARKWVPFKYKIDIEPRAPEFYFOQKIDYLDKDKVQPSFVKDRRAMKREYEEFKIRI 467
 Db 413 ETSEFARKWVPFKYKIDIEPRAPEFYFOQKIDYLDKDKVAASFVRERRAMKREYEEFKVRI 472
 QY 468 NALVSKALVPRREGTMDQGTWPNGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVVSR 527
 Db 473 NALVAKAQVPEGTMDQGSWPNGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVVSR 532
 QY 528 EKRPQGHKKKAGAMNALVRVSAVLINGQYMLNLDCDHYINNSKAVREAMCFMLDPLNLG 587
 Db 533 EKRPQGHKKKAGAMNALVRVSAVLNSAAYLLNLDCDHYINNSKAIKAMCFMMDPLVGK 592
 QY 588 OVCYVQFORFOGIDORNDRYANRNVFEDINLRGLDGTQGVYVYTCGVFNRTAIYGP 647
 Db 593 KVCYVQFORFOGIDORNDRYANRNVFEDINMKGLDGTQGVYVYTCGVFNRTAIYGP 652
 QY 648 PIKAKRP-----GFLASLCGGKKKASKKRSDDKKKS--NKHVDSVVPVFNLED 695
 Db 653 P-KTKKPPSKTNCWPKWLSCCCSRNKNKKTTPKTEKKKRLFFKKAENPSPAYALGE 711
 QY 696 IEBGVGAGFDDKSVLMSQMSLEKRFQGSAAFAVASTLMEYGGVQPSSTPESLKEAIVH 755
 Db 712 IDEGAPGA--DIEKAGIVNQOKLEKFKQSSVFAVASTLLENGGTLKSPASLKEAIVH 769
 QY 756 ISGCVEDKSEWTEIGWIYGSVTEIDLTGFKMHARGWSVCMKPRPAFKGSAPINLSDR 815
 Db 770 ISGCVEDKTDWCKEIGWIYGSVTEIDLTGFKMHARGWSVCMKPRPAFKGSAPINLSDR 829
 QY 816 LNOVLWALGSVEILFHSRCHPLWYGGYGRKFLERFAINTIYPLSLPLVYCIILPAI 875
 Db 830 LHOVLWALGSVEILFHSRCHPLWYGGYGRKFLERFINSIVPWTSIPLLAYCTIIPAI 889
 QY 876 CLITGXFIMPEISNLASLWIFALFLSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAH 935
 Db 890 CLITGXFITPELTNVASLWIFALFLSIFATGILEMRWSGVGIDEMWRNEQFWIGGISAH 949
 QY 936 LFAVFGLLKVLAGIDTNETVTSKANDEBGFDAELYMFKWTLLIPPTTILIIINWGVVA 995
 Db 950 LFAVFGLLKVLAGIDTNETVTSKANDEBGFDAELYMFKWTLLIPPTTILIIINWGVVA 1008
 QY 996 GTSYAINSGYQSGPLFGKLFPAFWIVHLXPLFKGLMGRONRPTIIVWAVLLASIFS 1055
 Db 1009 GISNALNNGYESWGLPFGKLFPAFWIVHLXPLFKGLMGRONRPTIIVWAVLLASIFS 1068
 QY 1056 LLWVRVDPETTRLAGNIOTCCINC 1080
 Db 1069 LLWVRVDPFLAKSNGPLLEECGLDC 1093

RESULT 11

Q9LLI4
 ID Q9LLI4 PRELIMINARY; PRT: 1059 AA.
 AC Q9LLI4
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cellulose synthase-6.
 GN CESA-6.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_Taxid=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20398328; PubMed=10938350;
 RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
 RA Xooconostle-Cazares B., Delmer D.P.;
 RT "A comparative analysis of the plant cellulose synthase (CesA) gene

RT family";
 RL Plant Physiol. 123:1313-1324 (2000).
 DR EMBL; AF200530; AAF99966.1; -.
 DR InterPro: IPR005150; Cellulose_synt.
 DR InterPro: IPR001841; Znf_ring.
 DR Pfam: PF03552; Cellulose_synt; 1.
 DR SMART: SM00184; RING; 1.
 SQ SEQUENCE 1059 AA; 120044 MW; 289DA26B25232249 CRC64;
 Query Match 66.1%; Score 3820; DB 10; Length 1059;
 Best Local Similarity 65.1%; Pred. No. 1.3e-300;
 Matches 714; Conservative 140; Mismatches 168; Indels 74; Gaps 16;
 QY 16 GDVQCICADGLGTLDGVDFTACDVCREFPVCRPCYEHKKEGTQACLOCKTKYKRRHGSP 75
 Db 6 GQVCQICGDDVGRNPDGEPFVACNECAPICDCYEEYRREGTQNCPOCKTRFRKFGCA 65
 QY 76 AIRBEGDDTDADGSDRNYPASGTEDQOKIADRMRSWRNMTGGSGVG-----HPKY 129
 Db 66 RVPGEDEEDGVDDLENEFNW---SDKHDSQYLAESMLHAHMSYGRGADLDGVPOFHP-- 120
 QY 130 DSGEIGLSKYDSGEIPRGVYPSVTNSOMSGEIPGASPDHMMSP---GNISRRAPPYV 186
 Db 121 -----IPN--VPLLTNGQWDDIP---PQHALVFSFVGGGGRKTHPLPYA 161
 QY 187 N-----HSPNPSREFSG-SIGNVAMKERVDMGWMKODKGAIPMNTNGTSIAPSEGRAAT 238
 Db 162 DPNLPVQPRSDPDKLAAYGYGVAMKERMESWKQER--MHQTR-----NDGGDD 213
 QY 239 DDASTEYNMEDALLNDTROPLSRKVPTASSKINPYBWIVLRLVLSIFLHYRLNPV 298
 Db 214 GDAA-----DLPLMDARQPLSRKIPLPSSQINPYRMIIILRLVLCFFHFRVHPV 266
 QY 299 RNAYPLWLVSICEIWFALSWILDQFPKWFPINRETYLDRLALRYDRBEGESQALAVDIF 358
 Db 267 PDFAFWLWLSVICEIWFAMSWILDQFPKWFPITERETYLDRLSLRFDKEGHPSQLAPVDF 326
 QY 359 VSTVDPLKPPVTANTVLSILAVDYPVDKVCYVSDGASMLTFDALAESEFARKWVP 418
 Db 327 VSTVDPLKPPVTANTVLSILSVDPYVDKVCYVSDGAAAMLFEALSESEFAKKWVP 386
 QY 419 FVKYDIEPRAPEFFYFOQKIDYLDKDKVQPSFVKDRRAMKREYEEFKIRINALVSKALKVP 478
 Db 387 FCKRYSLEPRAPEWYFOQKIDYLDKDKVAPFVRRERAMKREYEEFKVIRINALVAKAKVP 446
 QY 479 EGEWIMODGTWPNGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVYVSRERKPGFHHKK 538
 Db 447 EGEWIMODGTWPNGNTRDHPGMIQVFLHSGGLDTGEGNELPLRVYVSRERKPGYNNHKK 506
 QY 539 AGAMNALVRVSAVLINGQYMLNLDCDHYINNSKAVREAMCFMLDPLNPGVQVYVQFORF 598
 Db 507 AGAMNALVRVSAVLNAPYLLNLDCHDHYINNSKAIKAMCFMMDPLLGKKYVYVQFORF 566
 QY 599 DGIIDRDYANRNTVFFDINLRGLDGIQGVYVGTGCVFNRTAIYGPPTKAKKP---- 654
 Db 567 DGIIDRDYANRNVFFDINMKGLDGIQGVYVGTGCVFRQALYGYDAP-KTKKPPSRT 625
 QY 655 -----GFLASLCGGKKKASKKRSDDK-----KKSNNHVDSSVVPVFNLEDIEBVGAG 704
 Db 626 CNCWPKWICCCCGFNGRKTCKTKTKSPKFKIKKLFKKENQAPAYALGEIDEAAPGA- 684
 QY 705 FDDEKSVLMSQMSLEKRFQGSAAFAVASTLMEYGGVQPSSTPESLKEAIVHISGYEDKS 764
 Db 685 -ENKASIVNQOKLEKFKQSSVFAVASTLLENGGTLKSPASLKEAIVHISGYEDKT 743
 QY 765 EWGTEIGWYGSVTEIDLTGFKMHARGWSVCMKPRPAFKGSAPINLSRLNQLRWAL 824
 Db 744 GWKXIDIGWYGSVTEIDLTGFKMHARGWSVCMKPRPAFKGSAPINLSRDFHQLRWAL 803
 QY 825 GSVEILFHSRCHPLWYGGYGRKFLERFAINTIYPLSLPLVYCIILPAICTLTGTFEIM 884
 Db 804 GSVEILFHSRCHPLWYGGYGRKFLERFSTINSIVPWTSIPLLAYCTLPAICLTGTFEIT 863

```
QY 885 PEISLASIWFIALFLSIFATGILEMRWSGVGIDEMWNEQFWIGGISAHLFAVFOGLL 944
Db 864 PELNNVASLWFLSIFATGILEMRWSGVGIDEMWNEQFWIGGISAHLFAVFOGLL 923
QY 945 KVLADIDNFTVTSKANDEEGFAELMFKWTTLLIPPTTILIIINMGVAGTSYAINSG 1004
Db 924 KVIAGVDSFTVTSKGGDE-EFSELYTFKWTLLIPPTTLLNFIQVAGISAINNG 982
QY 1005 YSWGPLEGKFFFAFWIVHLVLPFLKGLMGQRNRTPTIVWAVLLASTFSLWVRDVF 1064
Db 983 YESWPLEGKFFFAFWIVHLVLPFLKGLMGQRNRTPTIVWAVLLASTFSLWVRDVF 1042
QY 1065 TTRLAGPNIQTGCING 1080
Db 1043 LAKDGPILLEGGGLDC 1058

RESULT 12
ID Q9LLI3 PRELIMINARY; PRT; 1086 AA.
AC Q9LLI3;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Cellulose synthase-7.
GN CESA-7.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398328; PubMed=10938350;
RA Holland N., Holland D., Heltentaris T., Dhugga K.S.,
RA Xoonostle-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (CesA) gene
family."
RL Plant Physiol. 123:1313-1324 (2000).
DR EMBL; AF200531; AAF89967.1; -.
DR InterPro; IPR001841; Znf.ring.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1086 AA; 122608 MW; DD03C73ABD13E2EB CRC64;

Query Match 66.1%; Score 3818; DB 10; Length 1086;
Best Local Similarity 64.7%; Pred. No. 1.9e-300;
Matches 716; Conservative 141; Mismatches 178; Indels 72; Gaps 17;

QY 2 DGDADALSGRAGDVQICADGLGTTLDGDVFTACDVCRFPVCRPCYEHKRGCTQAC 61
Db 23 DGDGPG-KPPREQNGQVCQICGDDVGLAPGDPFVACNECAFVCRDCYERREGTQNC 81
QY 62 LOCKTKYKRHRGSPALRGEGDDTDADGSDFNYPASGTEDQKTIADRMRSWRMNTGS 121
Db 82 PCKTRYKRLKQCRVGTGDEEDGVDDLDNEFNWDG----HDSQSVAESML----- 128
QY 122 GNVGHPKYDSGIGLSKYDSGEIPRGY-----VPSVTNSQMSGEIPGAPDHMMSP-- 174
Db 129 --YGHMSYGRGG-----DPNGAPQAFQLPNPNVPLLTNGQVDDIP---PEQHLYPSEF 177
QY 175 -GNISRRAPFPYVN-----HSPNPSREFSG-SIGNVAKKERVDCWKKQDKGAPMTN 225
Db 178 GGGGKRHLPLPADPSLPVQPRMSDPSKDLAAYGYGSVAKKERNKQQRER--MHQTG 235
QY 226 GTSIAPSEGRAATDIDASTEYNMEDALLNDETROPILSRKVPITASSKINPYMVIVLRYV 285
Db 236 -----NDGGDGDGDA-----DLPLMDEARQQLSRKIPLPSSQINPYMIIIRLVY 282
QY 286 LSIFLHYRTNPNRYAPLWLLSVCEIWFALSWILDQPKFPNPNRYLDRALRYDR 345
Db 283 LGFFFFHYRVMHPVNDALWLLSVCEIWFAMSWILDQPKFPNPNRYLDRALRYDR 342
```

RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
RT synthase required for secondary cell wall synthesis."
RL Plant Cell 11:769-780(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF091713; AAD40885.1; -
DR EMBL: AL391142; CAC01737.1; -
DR InterPro: IPR005150; Cellulose_synt.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synt; 1.
DR SMART: SM00184; RING; 1.
SQ SEQUENCE 1026 AA; 115797 MW; 503BFC78BB6E511 CRC64;

Query Match 55.8%; Score 3801; DB 10; Length 1026;
Best Local Similarity 67.8%; Pred. No. 4.2e-299;
Matches 725; Conservative 104; Mismatches 156; Indels 88; Gaps 17;

QY 16 GDVCOICADGLTLDGDFVTCACDVCPCYEHKRGKQACLOCKTKYKRRGSP 75
DB 16 GDVCOICADGLTLDGDFVTCACDVCPCYEHKRGKQACLOCKTKYKRRGSP 75

QY 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPCQKTRKLRGSP 93
DB 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPCQKTRKLRGSP 93

QY 76 AIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMNTGGSGNYPKYSGEIG 135
DB 76 AIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMNTGGSGNYPKYSGEIG 135

QY 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142
DB 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142

QY 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRAFPYVNHSPNP 192
DB 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRAFPYVNHSPNP 192

QY 143 -----NGRFPVPVIAGHSGEFPVGGYNGGEH-----GLHKRV-----HPYP 179
DB 143 -----NGRFPVPVIAGHSGEFPVGGYNGGEH-----GLHKRV-----HPYP 179

QY 193 SREFSGSIGNVANKERVGDGKMKODKGAIPMTNGTSIAPSEGRAADIDASTEYNMEDAL 252
DB 193 SREFSGSIGNVANKERVGDGKMKODKGAIPMTNGTSIAPSEGRAADIDASTEYNMEDAL 252

QY 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG----- 214
DB 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG----- 214

QY 253 LNDETROPLSRKPIASSKINPYRMVTLVRLVLSIFLHYRLTNVRYNAPYLLWSVICE 312
DB 253 LNDETROPLSRKPIASSKINPYRMVTLVRLVLSIFLHYRLTNVRYNAPYLLWSVICE 312

QY 215 LIDEAROLSRKPIASSKINPYRMVTLVRLVLSIFLHYRLTNVRYNAPYLLWSVICE 274
DB 215 LIDEAROLSRKPIASSKINPYRMVTLVRLVLSIFLHYRLTNVRYNAPYLLWSVICE 274

QY 313 IFWALSILQDFPKWPIERETYLDRLSLYEREGEPNMLAPVDVFSVVDLPKEPIYT 372
DB 313 IFWALSILQDFPKWPIERETYLDRLSLYEREGEPNMLAPVDVFSVVDLPKEPIYT 372

QY 275 IFWALSILQDFPKWPIERETYLDRLSLYEREGEPNMLAPVDVFSVVDLPKEPIYT 334
DB 275 IFWALSILQDFPKWPIERETYLDRLSLYEREGEPNMLAPVDVFSVVDLPKEPIYT 334

QY 373 ANTVLSILAVDYPVKVSCVSDGASMLTFDALAEISFARKWVPVKKYDIEPRAEF 432
DB 373 ANTVLSILAVDYPVKVSCVSDGASMLTFDALAEISFARKWVPVKKYDIEPRAEF 432

QY 335 SNTVLSILANDYPVEKISCVSDGASMLTFESLSETAFARKWVPVCKKFSIEPRAPEM 394
DB 335 SNTVLSILANDYPVEKISCVSDGASMLTFESLSETAFARKWVPVCKKFSIEPRAPEM 394

QY 433 YFCQKIDYLDKQVPSVKDRRAKREYEFKIRINALYSKALKVPEEGWIMODGTPWPG 492
DB 433 YFCQKIDYLDKQVPSVKDRRAKREYEFKIRINALYSKALKVPEEGWIMODGTPWPG 492

QY 395 YFLKVDYLDQKVHPFVKERRAKREYEFKIRINAKVASKVPLEGIMODGTPWPG 454
DB 395 YFLKVDYLDQKVHPFVKERRAKREYEFKIRINAKVASKVPLEGIMODGTPWPG 454

QY 493 NNTRDHPGMIQVFLGHSGGLDTGEGNELPLRVLYVYSREKRPFGQHHKAGAMNALVRVSAL 552
DB 493 NNTRDHPGMIQVFLGHSGGLDTGEGNELPLRVLYVYSREKRPFGQHHKAGAMNALVRVSAL 552

QY 455 NNTKDHPGMIQVFLGHSGGDFVGEHELPLRVLYVYSREKRPFGQHHKAGAMNALVRVAGYL 514
DB 455 NNTKDHPGMIQVFLGHSGGDFVGEHELPLRVLYVYSREKRPFGQHHKAGAMNALVRVAGYL 514

QY 553 TNGOYMLNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 612
DB 553 TNGOYMLNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 612

QY 515 TNAPFMNLNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 574
DB 515 TNAPFMNLNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 574

QY 613 VFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 668
DB 613 VFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 668

QY 575 VFDINMKGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 632
DB 575 VFDINMKGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 632

QY 669 KSKRRSDKKHKNHVDSSVPVFNLEDIEEGVAGAGFDEKSVMSQMSLEKRFQGSAAF 728
DB 669 KSKRRSDKKHKNHVDSSVPVFNLEDIEEGVAGAGFDEKSVMSQMSLEKRFQGSAAF 728

QY 633 -----NKKFSKNDMNGDVAALG-----GABG-----DKEHLMSEMFEKTFGQSSIF 674
DB 633 -----NKKFSKNDMNGDVAALG-----GABG-----DKEHLMSEMFEKTFGQSSIF 674

QY 729 VASTLMEYGVGPOSSPESILLKKAHIVISCGYEDKSEWGTGTEIGWIYGSVTEDILTGFKMH 788
DB 729 VASTLMEYGVGPOSSPESILLKKAHIVISCGYEDKSEWGTGTEIGWIYGSVTEDILTGFKMH 788

DB 675 VTSTLMEEGVPPSSSPAVLLKKAHIVISCGYEDKTEGTGLWIYGSITEDILTGFKMH 734
QY 789 ARGWRSVYCMKPRPAFKGSAFINLSDRNLQVLRWALGSVEILFSRHCPPLWGY--GGRKLF 847
DB 735 CRGWSIYCMKPRPAFKGSAFINLSDRNLQVLRWALGSVEILFSRHCPPLWGYGKGLKW 794
QY 848 LERFAYINTTYPTSLPLLYCICLPAICLLTGKFIPEISNLASIFWALFLSIFATGI 907
DB 795 LERFAYANTTYPTSLPLLYCICLPAICLLTGKFIPEISNLASIFWALFLSIFATGI 854
QY 908 LEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFOGLLKLVLGIDTNTVTSKANDESGDF 967
DB 855 LELRWSGVGIDEMWRNEQFWIGGISAHLFAVFOGLLKLVLGIDTNTVTSKANDESGDF 913
QY 968 AELYMFKWTTLLIPPTTILINMVGWVAGTSYAINSGYQSGWGLFGKLFPAFWIVHLYP 1027
DB 914 GELYAFKWTLLIPPTTILINMVGWVAGTSYAINSGYQSGWGLFGKLFPAFWIVHLYP 973
QY 1028 FLKGLMGRQNRTPPIVIVWAVLLASIFSLWVRVDPFTTTRLAGNPQTGCGINC 1080
DB 974 FLKGLMGRQNRTPPIVIVWAVLLASIFSLWVRVDPFTTTRLAGNPQTGCGINC 1026

RESULT 14
Q9XHP6 PRELIMINARY; PRT; 1026 AA.
ID Q9XHP6;
AC Q9XHP6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN IRX3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
ON NCBI_taxid=3702;
RX NCBI_SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG RECTA;
RX MEDLINE=99264300; PubMed=10330464;
RA Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
RL synthase required for secondary cell wall synthesis.";
DR EMBL: AF088917; AAD32031.1; -
DR InterPro: IPR005150; Cellulose_synt.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synt; 1.
DR SMART: SM00184; RING; 1.
SQ SEQUENCE 1026 AA; 115858 MW; 453BFD1D283CAD70 CRC64;

Query Match 65.7%; Score 3795; DB 10; Length 1026;
Best Local Similarity 67.5%; Pred. No. 1.3e-298;
Matches 724; Conservative 104; Mismatches 157; Indels 88; Gaps 17;

QY 16 GDVCOICADGLTLDGDFVTCACDVCPCYEHKRGKQACLOCKTKYKRRGSP 75
DB 16 GDVCOICADGLTLDGDFVTCACDVCPCYEHKRGKQACLOCKTKYKRRGSP 75

QY 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPCQKTRKLRGSP 93
DB 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPCQKTRKLRGSP 93

QY 76 AIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMNTGGSGNYPKYSGEIG 135
DB 76 AIRGEGDDTDADGSDFNYPASGTEDQKQIADRMRSWRMNTGGSGNYPKYSGEIG 135

QY 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142
DB 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142

QY 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRAFPYVNHSPNP 192
DB 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRAFPYVNHSPNP 192

QY 143 -----NGRFPVPVIAGHSGEFPVGGYNGGEH-----GLHKRV-----HPYP 179
DB 143 -----NGRFPVPVIAGHSGEFPVGGYNGGEH-----GLHKRV-----HPYP 179

QY 193 SREFSGSIGNVANKERVGDGKMKODKGAIPMTNGTSIAPSEGRAATDIDASTEYNMEDAL 252
DB 193 SREFSGSIGNVANKERVGDGKMKODKGAIPMTNGTSIAPSEGRAATDIDASTEYNMEDAL 252

QY 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG----- 214
DB 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG----- 214

QY 253 LNDETROPLSRKVPPIASSKINPYRMVIVLRLVLSIFLHYRLNPNVRNAYPLMLLSVICE 312
Db 215 LIDEARQPLSRKVPPIASSKINPYRMVIVLRLVLSIFLHYRLNPNVRNAYPLMLLSVICE 274
QY 313 INFALSWILDQFPKWPPIINRETYDLRLALRYDREGSPSQAALAVDIFVSTVDPLKEPIVT 372
Db 275 INFALSWILDQFPKWPPIINRETYDLRLALRYDREGSPSQAALAVDIFVSTVDPLKEPIVT 334
QY 373 ANTVLSILAVDYPVDKVCVSDGASMLTFDALAETSEFARKWVPVKKYDTEPRAPEF 432
Db 335 SNTVLSILAVDYPVDKVCVSDGASMLTFDALAETSEFARKWVPVKKYDTEPRAPEM 394
QY 433 YFCQKIDYLDKVPQFVDRRRAMKREYEEFKIRINALVSKALKVPBEGWIMQDGPWPG 492
Db 395 YFTLKVDYLDKVPQFVDRRRAMKREYEEFKIRINALVSKALKVPBEGWIMQDGPWPG 454
QY 493 NNTFRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSRKRGFGHKKAGAMNALVRVSAVL 552
Db 455 NNTFRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSRKRGFGHKKAGAMNALVRVSAVL 514
QY 553 TNGOYMLNLDCHYINNSKAVREAMCFMDPNLGPQVYVQFQFQFQFQFQFQFQFQFQFQF 612
Db 515 TNAPFMNLDCHYINNSKAVREAMCFMDPNLGPQVYVQFQFQFQFQFQFQFQFQFQFQF 574
QY 613 VFEDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPPKAKPGFLASLC----GGRKKAS 668
Db 575 VFEDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPPKAKPGFLASLC----GGRKKAS 632
QY 669 KSKRSDDKKSNKHVDSSVYPNLEDEIEGVGEGAFDEDEKSVLMSOMSLKRFQGSAAF 728
Db 633 KSKRSDDKKSNKHVDSSVYPNLEDEIEGVGEGAFDEDEKSVLMSOMSLKRFQGSAAF 674
QY 729 VASTLMEYGVGPOSSPESILLKAIHVISGVEDKSEWTEIGWIGYSVTEIDILTGFKM 788
Db 675 VASTLMEYGVGPOSSPESILLKAIHVISGVEDKSEWTEIGWIGYSVTEIDILTGFKM 734
QY 789 ARGWRSYCMKPKPAFGKSGAPINLSRDLNOVLWALGVSVEILSRHCPWLYGY--GGRKLF 847
Db 735 CRGWSYCMKPKPAFGKSGAPINLSRDLNOVLWALGVSVEILSRHCPWLYGY--GGRKLF 794
QY 848 LERFAYINTIYPLTSLPLVAYCILPAICLLTGKFTMPETSINLASIWFALFSLFATGI 907
Db 795 LERFAYINTIYPLTSLPLVAYCILPAICLLTGKFTMPETSINLASIWFALFSLFATGI 854
QY 908 LEMRWGSGVDEWRNEQFWVIGGISAHLFAVFGGLKVLGAGIDTNTFTVTSKANDERGF 967
Db 855 LEMRWGSGVDEWRNEQFWVIGGISAHLFAVFGGLKVLGAGIDTNTFTVTSKANDERGF 913
QY 968 AELYMFKWTTLLPPTTILLINMVGVVAGTSYAINSGYQSWGDLFGKLFPAFVIVHLYP 1027
Db 914 GELYAFKWTLLPPTTILLINMVGVVAGTSYAINSGYQSWGDLFGKLFPAFVIVHLYP 973
QY 1028 FLKGLMGRQNTTIVVAVLASIFSLWVRVDPPTTLRAGFNIGTCGINC 1080
Db 974 FLKGLMGRQNTTIVVAVLASIFSLWVRVDPPTTLRAGFNIGTCGINC 1026

RESULT 15
Q9SKJ5 PRELIMINARY; PRT; 1065 AA.
AC Q9SKJ5;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Putative cellulose synthase catalytic subunit.
GN AT2G25540.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10
RA Lin X., Kaul S., Rounsley S.
RA Fujii C.Y., Mason T.M., Bow
RA Buell C.R., Ketchum K.A.,
RA Cronin L.A., Shen M., Van
RA Adams M.D., Carrera A.J., C
RA Copenhaver G.P., Preuss D.
RA Salzberg S.L., Fraser C.M.,
RT "Sequence and analysis of C
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the
DR EMBL; AC006300; AAD20713.1;
DR InterPro; IPR005150; Cellul
DR InterPro; IPR001309; ICE_P2
DR InterPro; IPR001841; Znf_r
DR Pfam; PF03552; Cellulose_sy
DR SMART; SM00184; RING; l.
DR PROSITE; PS01121; CASPASE
SQ SEQUENCE 1065 AA; 120618
Query Match 65.4%
Best Local Similarity 64.1%
Matches 701; Conservative 1
QY 2 DGADALSKGRHGAGDVQCICA
Db 16 DSDGDLKPLKDLNGQICQIC
QY 62 LQCKTKYKRGSPAIRGEGD
Db 76 PCKARFRNGSPRVEDEKE
QY 122 GNV-----GHPKYDSGEIG
Db 133 ESLPSVLLTHGHP-----
QY 174 TGNISRRAPPY-----VNHS
Db 162 C--IDPLPGIYQLLLPVRI
QY 228 SIAPSEGAATDIDASTENNE
Db 218 --KYHEKGGE--PEGTSGNG
QY 288 IFLHYRLNPNVRNAYPLMLLSV
Db 274 VELHYRTHPVKDAYALWLTSL
QY 348 EPSOLAAYDIFVSTVDPLKEPE
Db 334 EPSOLAAYDIFVSTVDPLKEPE
QY 408 ETSEFARKWVPVKKYDIEPRA
Db 394 ETAEFSKKWVPFCKKFNIEPR
QY 468 NALVSKALKVPBEGWIMQDGT
Db 454 NILVAKAKIPEDEGTWEDGTS
QY 528 EKRPFGHKKAGAMNALVRVS
Db 514 EKRPFGHKKAGAMNALVRVS
QY 588 QVCYVQFPQRFQDIDRNDRYAN
Db 574 KCCYVQFPQRFQDIDRNDRYAN

97; Shea T.P., Benito M.-I., Town C.D.,
S.L., Barnstead M.E., Feldblum T.V.,
J., Ronning C.M., Koo H., Moffat K.S.,
E., Umayam L., Tallon L.J., Gill J.E.,
y T.H., Goodman H.M., Somerville C.R.,
man W.C., White O., Eisen J.A.,
er J.C.;
osome 2 of the plant Arabidopsis
GenBank/DBJ databases.
ynt.
ore 3778.5; DB 10; Length 1065;
ad. No. 3e-297;
Mismatches 172; Indels 69; Gaps 14;
TTLDGDFVTAQDVCRCVPCYHEREGTQAC 61
LTKTGNFVACNECGFPLCQCYEYERDGSQC 75
DDGSDFNYPASSTEDOKOKIADRMRSWRMTGGS 121
DIENEFDTQG---NNKARLPHRAEEFSSSSRHE 132
SGEIPRGVVPVSTNSQMSGEIPGASPDHH--WMSP 173
-----VSGEIP--TPDRNATLSP 161
SF--SGSIGNVAWKEBVDGKWKQDKGATPMTNGT 227
DLNSYGLVNDKMKRIQCKMLKQDKNMLHMTG-- 217
NDETROPLSRKVPPIASSKINPYRMVIVLRLVLS 287
DDARLPMSRVVHPSPASRTYRIVIVLRLIILG 273
AFALSWILDQFPKWPPIINRETYDLRLALRYDREG 347
AFALSWILDQFPKWPPIINRETYDLRLALRYDREG 333
NTVLSILAVDYPVDKVCVSDGASMLTFDALA 407
NTVLSILAVDYPVDKVCVSDGASMLTFDALA 393
FCQKIDYLDKVPQFVDRRRAMKREYEEFKIRI 467
FSQKIDYLDKVPQFVDRRRAMKREYEEFKIRI 453
NTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSR 527
NTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVSR 513
NGOYMLNLDCHYINNSKAVREAMCFMDPNLGP 587
NGOYMLNLDCHYINNSKAVREAMCFMDPNLGP 573
FFDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPP 647
FFDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPP 633

```
QY 648 PIKAK--KPGFLASLCGGKKKASKK RSDKKKSNKHVDSSVPVFNLEDIEGVEGAG 704
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
634 VLTEEDLENIIVKSCFGSKKSKRKIPNYEDNRSIKRSDSNVFLFNMEDIDEDVE--G 691
QY 705 FDEKSVLMSQMSLEKRCOSAAFVASTLMEYGGVQSQSTPESLLKEAIIHVISCYEDKS 764
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
692 YEDMSLLVSQRLEKRFQSPVFAATFMEOGGLPSTNPLTLLEAIIHVISCYEAKT 751
QY 765 EMGTEIGWIYGSVTEIILTGFKNHARGWRSVYCMKRPAPKGSAPINLSDRLNQVLRWAL 824
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
752 DAGKEIGWIYGSVTEIILTGFKNHARGWISYCVPSRPAFKGSAPINLSDRLNQVLRWAL 811
QY 825 GSVEILFSRCHPLWYGYGGRLEKFLERFAYINTIYPLTSLPLLYCYILPAICLTGKFIM 884
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
812 GSTEILLSRHCPWIYGYNGRLKLERIAYINTIVYPTSIPLLAYCMLPAFCLITNTFII 871
QY 885 PEISNLASTWFIATFLSIFATGILEMRWSGVGIDEMWRNEOFWVIGGISAHLEAVFOGLL 944
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
872 PEISNLASLCFLLFASIYASALLEKWSDALEDDWRNEQFWVIGYSAHLEAVFOGLL 931
QY 945 KVLAGIDTNTVTSKANDEGDFAEIYMKWTLLIPPTTILINMGVWAGTSYAINSG 1004
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
932 KVFAgidTNTVTSKASDEGDFAEIYVFKWTSLLIPTTILLYNLVGIAGVSYAINSG 991
QY 1005 YQSWGPLFGKLFPAFWVIHLYPFLKGLMGQRNRTPTIVVWVLLASIFSLWVRVDPF 1064
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
992 YQSWGPLMGKLLFAFWVAHLYPFLKGLGRNRTPTIVVWSALLASIFSLWVRINPF 1051
QY 1065 TTRLAGPNIOTCGI 1078
Db : : : : :
1052 -----VSTTGV 1057
```

Search completed: February 19, 2003, 16:36:55
Job time : 63 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: February 19, 2003, 16:34:12 ; Search time 25 Seconds
(without alignments)
1271.070 Million cell updates/sec

Title: US-09-900-237-30

Perfect score: 5778

Sequence: 1 MDGADALKSRHGAGDVQ.....VDPFTRLAGPNQTCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4455.5	77.1	881	4	US-08-960-048-8
2	3477	60.2	974	4	US-08-960-048-6
3	2811	48.7	685	4	US-08-960-048-7
4	268	4.6	693	4	US-08-960-048-11
5	255.5	4.4	861	4	US-08-960-048-12
6	238.5	4.1	3031	1	US-07-689-008-2
7	237.5	4.1	756	4	US-08-960-048-10
8	237.5	4.1	756	4	US-08-960-048-9
9	207	3.6	723	4	US-08-960-048-9
10	207	3.6	723	5	PCT-US91-01726-4
11	114	2.0	357	1	US-08-119-773-4
12	112	1.9	1957	4	US-08-669-656A-8
13	111	1.9	1957	4	US-08-669-656A-2
14	108	1.9	1167	1	US-08-485-568A-6
15	108	1.9	1167	2	US-08-590-554A-6
16	108	1.9	1167	2	US-08-184-223-6
17	108	1.9	1168	1	US-08-620-717A-9
18	107.5	1.9	2132	4	US-08-669-656A-6
19	106	1.8	357	1	US-08-119-773-6
20	105	1.8	357	1	US-08-119-773-5
21	105	1.8	357	1	US-08-119-773-2
22	104.5	1.8	3033	1	US-07-925-695-5
23	102	1.8	616	4	US-09-136-574A-47
24	101	1.7	418	4	US-09-134-001C-4051
25	101	1.7	1026	1	US-07-998-003A-95
26	101	1.7	1026	1	US-08-453-274B-95
27	101	1.7	1026	1	US-08-453-695A-95

28	101	1.7	1026	1	US-08-268-161A-95	Sequence 95, Appl
29	101	1.7	1026	1	US-08-453-702A-95	Sequence 95, Appl
30	101	1.7	1026	4	US-09-099-639-95	Sequence 95, Appl
31	101	1.7	1026	5	PCT-US93-12588-95	Sequence 95, Appl
32	101	1.7	1026	5	PCT-US95-08071-95	Sequence 95, Appl
33	101	1.7	1203	1	US-07-998-003A-103	Sequence 103, App
34	101	1.7	1203	1	US-08-453-274B-103	Sequence 103, App
35	101	1.7	1203	1	US-08-453-695A-103	Sequence 103, App
36	101	1.7	1203	1	US-08-268-161A-103	Sequence 103, App
37	101	1.7	1203	2	US-08-453-702A-103	Sequence 103, App
38	101	1.7	1203	4	US-09-099-639-103	Sequence 103, App
39	101	1.7	1203	5	PCT-US93-12588-103	Sequence 103, App
40	101	1.7	1203	5	PCT-US95-08071-103	Sequence 103, App
41	100.5	1.7	1956	4	US-08-843-417-2	Sequence 2, Appl
42	99.5	1.7	2206	1	US-07-852-260-2	Sequence 2, Appl
43	99.5	1.7	2206	1	US-08-461-503-2	Sequence 2, Appl
44	99.5	1.7	2206	4	US-08-465-250-2	Sequence 2, Appl
45	98.5	1.7	3033	1	US-07-925-695-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-08-960-048-8
; Sequence 8, Application US/08960048C

; Patent No. 6271443

; GENERAL INFORMATION:

; APPLICANT: Stalker, D. et al.

; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter

; FILE OF INVENTION: Sequences

; FILE REFERENCE: 15621/01/US

; CURRENT APPLICATION NUMBER: US/08/960,048C

; PRIOR FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987

; PRIOR FILING DATE: 1996-10-29

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 881

; TYPE: PRT

; ORGANISM: Oryzae sativa

US-08-960-048-8

Query Match 77.1%; Score 4455.5; DB 4; Length 881;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;

QY 201 GNVAWKEVDGKWKQDKGAIPMTNGTSIAPSEGRAATDIDASTEXNMEDALLNDETROP 260

Db 1 GNVAWKEVDGKWKQDKGAIPMTNGTSIAPSEGRGVGDIDASTDYNNEDALLNDETROP 60

QY 261 LSRKVPDIASSKINPYRMVIVLVLVLSIFLHYRLTNPRNAYPLWLLSVICETWFLSWI 320

Db 61 LSRKVPDIASSKINPYRMVIVLVLVLSIFLHYRLTNPRNAYPLWLLSVICETWFLSWI 120

QY 321 LDQFPKPPINRETYLDRLALRYDREGPEPSOLAAYDIFVSTVDPLKEPPIVTTANTVLSIL 380

Db 121 LDQFPKPPINRETYLDRLALRYDREGPEPSOLAAYDIFVSTVDPLKEPPIVTTANTVLSIL 180

QY 381 AVDYVDKVCYVSDDGASMLTFDALAETSEFARKWPFVKYKIDIEPRAPEFFCOKIDY 440

Db 181 AVDYVDKVCYVSDDGASMLTFDALAETSEFARKWPFVKYKIDIEPRAPEFFCOKIDY 240

QY 441 LKDKVQPSFKDRRAMKREYEERKIRINALVSKALVPEEGWTMQDGTWPGNNTDRHPG 500

Db 241 LKDKVQPSFKDRRAMKREYEERKIRINALVSKALVPEEGWTMQDGTWPGNNTDRHPG 300

QY 501 MIQVFLHSGGLDTEGNELPRLVYSREKRPFGQHHKAGAMNALVRVSAVLNNGOYMLN 560

Db 301 MIQVFLHSGGLDTEGNELPRLVYSREKRPFGQHHKAGAMNALVRVSAVLNNGOYMLN 360

QY 561 LDCDHYINNSKAVREAMCFLMDPNLGPQVCYVQFPQFDGIDRNDRYANRNTVFFDINLR 620

Db 361 LDCDHINNKAALREACFLMDNLSRVSYYQVFPQRFQDIDRNDRYANRNTVFEDINUR 420
QY 621 GLDGIQGVYVGGVCFNRTAIYGYEPPKAKKPG-FLASLCGGKKKASKKRSDDKK 679
Db 421 GLDGIQGVYVGGVCFNRTAIYGYEPPKAKKPGFLASLCGGKKKASKKRSDDKK 480
QY 680 SNKHVDSSVPVFNLEDIEEGVEGAGFDDKSKVLSQMSLEKRFQSGAFAVASTLMYEGGV 739
Db 481 SNKHVDSSVPVFNLEDIEEGVEGAGFDDKSKVLSQMSLEKRFQSGAFAVASTLMYEGGV 540
QY 740 POSTPESLLEKRAHVISGCGYEDKSEMGTEIGWIGSVTEIDLTGFKMHARGRSVYCNP 799
Db 541 POSTPESLLEKRAHVISGCGYEDKSEMGTEIGWIGSVTEIDLTGFKMHARGRSVYCNP 600
QY 800 KRPAFGSAPINLSRNLQVLRWALGSVELFSRHCPLWYGGRLKFLERFAYINTTII 859
Db 601 KRPAFGSAPINLSRNLQVLRWALGSVELFSRHCPLWYGGRLKFLERFAYINTTII 660
QY 860 PLTSLPVLVYICILPAICLTGTFIMPEISNLSINFIALFSLFATGILEMRWSGVGIDE 919
Db 661 PLTSLPVLVYICILPAICLTGTFIMPEISNLSINFIALFSLFATGILEMRWSGVGIDE 720
QY 920 WRNEQFWYIGTSAHLFAVFGQLKVLZAGIDTNTFTVTSKANDDEGDFAEIYMKWTLL 979
Db 721 WRNEQFWYIGTSAHLFAVFGQLKVLZAGIDTNTFTVTSKANDDEGDFAEIYMKWTLL 780
QY 980 IPTTILINMGVAVAGTSYAINSGYQSWGPLFGKLFPAFAFVHLYPLKGLMGQRNRT 1039
Db 781 IPTTILINMGVAVAGTSYAINSGYQSWGPLFGKLFPAFAFVHLYPLKGLMGQRNRT 840
QY 1040 PTIVVAVILLASIFSLLMVRVDPFTTRLAGPNIQTCGINC 1080
Db 841 PTIVVAVILLASIFSLLMVRVDPFTTRVTPGTPTQTCGINC 881

RESULT 2
US-08-960-048-6
; Sequence 6, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: Sequences
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-08-960-048-6
Query Match 60.2%; Score 3477; DB 4; Length 974;
Best Local Similarity 61.5%; Pred No. 1.9e-318;
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;

QY 18 VCOICADGLGTLDDGVFTACDVCRCPCYEHKRGTOACQCKTKYKRHRGSPA 77
Db 8 VCHTCGERHVLNVDGEPFVACHCECNFPICKSCFEYDLKGRKACLCR-----GSPY- 58
QY 78 RGEEDGDDADGSDNFYASCTEDQKQKIDMRMRWMTGSGNVG-HPKYDSGEIGL 136
Db 59 -----DENLDD-----VEKATGQOSTMAA-----HLNKSQDVGIHARHIS---SV 96
QY 137 SKYDSGEIPRGVYVTSNQSCEIPGASPDHMMSPCTGNISRRAPPFYVNHSPNPSREF 196
Db 97 STLDS-----EM 103

QY 197 SSGIGNVAVKERVDCGKMKQDK 256
Db 104 AEDNGNSIWKNRVESWKEKNK 154
QY 257 TRQPLSRKVPATASSKINPYRMV 316
Db 155 S-OPLSTIIPKPSRLAPYRTV 213
QY 317 LSWILDQPKKWPFPNRETYLDR 376
Db 214 FSWILDQPKKWPFPNRETYLDR 273
QY 377 LSILAVDPYDKVSCYVSDDG 436
Db 274 LSILALDYPYDKVSCYVSDDG 333
QY 437 KIDYLDKQVQSFVKDRRAMKR 496
Db 334 KIDYLDKQVQSFVKERRAMKR 393
QY 497 DHPGMIQVFLGSHGGDLTEGNE 556
Db 394 DHPGMIQVFLGSGARDIEGNE 453
QY 557 YMLNLDCHYINNSKAVREANG 616
Db 454 FILNLDCHYVNNKAVREANG 513
QY 617 INLRGLDGIQGVYVVGTCVFN 676
Db 514 VNMGLDGIQGVYVVGTCVFN 573
QY 677 KKSNNKHVDSSVPVFNLEDIE 735
Db 574 LYRDAKREELDAALFNLEID 627
QY 736 YGVVQPSPTPESLLEKRAHVIS 795
Db 628 NGVAESANPSTLKEAIVHIS 687
QY 796 YCMPKRPAPKGSAPINLSDR 854
Db 688 YCMPLRPAKGSAPINLSDR 747
QY 855 NTTYPLTSLPLLYVCILPAI 914
Db 748 NTIYPTFTSLPIAYCPLAI 807
QY 915 VGIDWWRNEQFWIGGISAH 974
Db 808 VSIEDLWRNEQFWIGVSAH 866
QY 975 WTTLLIPPTTLLIINMGVVA 1034
Db 867 WTTLLIPPTTLLIINMGVVA 926
QY 1035 RQNRPTIIVVAVILLASIFS 1080
Db 927 RQNRPTIIVLWSVILLASIFS 974

RESULT 3
US-08-960-048-7
; Sequence 7, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: Sequences
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 6271443

; PRIOR FILLING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-08-960-048-7

Query Match 48.7%; Score 2811; DB 4; Length 685;
Best Local Similarity 74.1%; Pred. No. 6.1e-256;
Matches 509; Conservative 78; Mismatches 76; Indels 24; Gaps 6;
QY 413 ARKWPFEVKKYDIEPRAPFEYFCOKIDYLDKQVQSFYKDRAMKREYEEFKIRINALVS 472
DB 1 ARRWPFCKKHNVPRAPFEYNEKIDYLDKQVHPSFKERRAMREYEEFKIRINALVA 60
QY 473 KALKVPEEGWIMODGTPWGNTRDHPGMIQVFLGHSGLDTEGNEPLRLVVSREKRP 532
DB 61 KAQKKEEGWVMDGTPWGNTRDHPGMIQVYLSAGALVDGKELPRLVVSREKRP 120
QY 533 FQHHKAGAMNALVRVSALVTNGQYMLNDCDHYINNSKAVREAMCFMDPDLGQVQCYV 592
DB 121 YQHHKAGAMNALVRVSALVTNAPFLNDCDHYINNSKAMREAMCFMDPDLGQVQCYV 180
QY 593 QPQRTDGDINRDYRANRNVTFEDINLRGLDGIQGVVYGTGCVENRTAIYGEPIKAK 652
DB 181 QPQRTDGDINRDYRANRNVTFEDINLRGLDGIQGVVYGTGCVENRTAIYGEPIKAK 240
QY 653 K-----PGFLASLCGKKKASKSKRS-----SDKK-----SNKHVDSSVPVFN 692
DB 241 RPKMTDCWPSCWCCCGGSRKSKKKEKGLGLGLLYGKKKMMGNKYKKSAPVFD 300
QY 693 LEDIEEGVEGAGDD--EKSVLMSQMSLEKRFQGSAAFAVASTIMEYGGVQPSSTPESLKE 751
DB 301 LEEIEEGLE--GYEELEKSTLMSQKNEKRFQGSPPVFIASITLMENGGLPEGTNSTSLIKE 358
QY 752 AIHVISCGYBDSKSEWTEIGWYGSVTEIDLTGFKMHARGWSVCMKRPAPKGSAPIN 811
DB 359 AIHVISCGYEKTEWKEIGWYGSVTEIDLTGFKMHARGWSVCMKRPAPKGSAPIN 418
QY 812 LSRLNQLRWALSGVEILFSRHCPHLYGGRKFLERFAFYNTIYPLTSLPLLYYCI 871
DB 419 LSRLHQLRWALSGVEILFSRHCPHLYGGRKFLERFAFYNTIYPLTSLPLLYYCI 478
QY 872 LPAICLLTGKFIIMPEISNLASIFALFSLIFATGILEMRWSGVGIDEMWNEQFWVIGG 931
DB 479 IPAVCLLTGKFIIPTLSNTSVWFLALFSLIATGLEDLWSGVSIQDWNRNEQFWVIGG 538
QY 932 ISAHFLFAVQGLLKLVLGIDTNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTILIINMV 991
DB 539 VSAHLFAVQGLLKLVLGIDTNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTILIINMV 597
QY 992 GVYAGTSYALNSGYSQSGPLGKLFPAFWVIVHLFPFLKGLMGRQNRPTIIVWAVLLA 1051
DB 598 GVYAGVSDALNNGYSQSGPLGKLFPAFWVIVHLFPFLKGLMGRQNRPTIIVWAVLLA 657
QY 1052 SIFSLWVRVDPPTTIRLAGPNIQTCGI 1078
DB 658 SIFSLWVRVDPPTTIRLAGPNIQTCGI 684

RESULT 4
US-08-960-048-11
; Sequence 11, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; CURRENT FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-960-048-11

Query Match 4.6%; Score 268; DB 4; Length 693;
Best Local Similarity 19.5%; Pred. No. 4.1e-16;
Matches 156; Conservative 83; Mismatches 201; Indels 360; Gaps 29;
QY 263 RKVPASIKNPYRMVIVLRVLVSIFLHYRLT-----NPVNAYPLMLLSVICBIWA 316
DB 11 RRMPPRESAL-----MLIVLSLTVSCRYIMWRYTSTLNWDDPVSLSGLLILFAITVAV 66
QY 317 LSWILDDOPKWFPIINRETYLDRLALRYDRGEPSSQLAAVDIFVSTYVDPLKEPIVTANTV 376
DB 67 L--VLGYFQVYVPLNRQP-----VPLPKDMSLWPS-----VDIFVPTYN-----EDLNWVKNTI 113
QY 377 LSLIADVPYDKVSCYVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPEYFCQ 436
DB 114 YASIGIDMPKDKLNILWDDG-----135
QY 437 KIDYLDKQVPSFKDRRAMKREYEEFKIRINALVSKALKVPEEGWIMODGTPWPGNNTR 496
DB 136 -----REEFRQAQNVG-----147
QY 497 DHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKRPQFQHHKAGAMNALVRVSALVTNGQ 556
DB 148 -----VKYIARTT-----HEHAKAGNINNAKYA-----KGE 174
QY 557 YMLNDCDHYINNSKAVREAMCFMDPDLGQVQCYVQFQRF-----DGIDRN-----DRYAN 609
DB 175 FVSIFDCDHYVTPRSFLQMTMGWELKE-----KOLAMQTPHHFSPDPFERNLGRFRKTPN 230
QY 610 RNTVFEDINLRGLDGIQGVVYGTGCVENRTAIYGEPIKAKKPGFLASLCGKKKASK 669
DB 231 EGTUFLYGLVODGNDMDATFFCGSCAVIR-----Kp-----262
QY 670 SKRSSDKKSNKHVDSSVPVFNLEDIEGVEGAGDFDEKSVLMSQMSLEKRFQGSAAFV 729
DB 263 -----262
QY 730 ASTIMEYGGVQPSSTPESLKEAHLVISCYEDKSEWTEIGWYGSVTEIDLTGFKMHA 789
DB 263 ---LDEIGGI-----AVE-----TVTDAHTSURLHR 286
QY 790 RGMRSVCMKRPAPKGSAPINLSDRNLQVLRWALSGVEILFSRHCPHLYGGRKFLER 849
DB 287 RGYTSAYM--RIPQAAGLATESLSAHIGQRIWARGVQI--FRLDNPL---TGKGLKFAQ 340
QY 850 RFAYINTIYPLTSLPLLYYCIIP-----AICLLTGKFIIM-----885
DB 341 RLCYVNAFMHFLSGIPRLIFLTAPLFLHLLHAYIYAPALMIALFVLPHHIHASLTNSKI 400
QY 886 -----EISNLASIFIA-----LFLSIFATGILEMRWSGVGIDEMWNEQFWV 929
DB 401 QGKYRHSFWEIYETVLAWIAPTLVALLNPHKGFENVTAKGGGLVE-----EYVDWVI 456
QY 930 GGISAHFLFAVQGLLKLVLGIDTNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTIL--I 987
DB 457 S--RPYIFLVLNLLVGVAVGI-----WRYFYGPPTTEMLTVV 490
QY 988 INWVGV-----VAGTSYAIN 1002
DB 491 VSMWVYFNILVLGGAVAVS 510

RESULT 5
US-08-960-048-12

```

NUCLEIC ACID SEQUENCES FOR THE
OF CELLULOSE SYNTHASE OPERON

Brown & Enersen
ter

OS
d. No. 3.1e-12;
,008

4
6

re 238.5; DB 1; Length 3031;
d. No. 3.1e-12;
Mismatches 232; Indels 335;
LLSVICEIWFALSMILDO---FPKWPP---
:| |: : | |: | |: | |: | |: |
EVL$ALV$CRY-LTWRLTETLDFDTWIOGGLGS
-----AVDIFVSTVDPLKEPPIVTA
LPLLPDNVDDMPTVDIFIPTVD---EQLSIVR
FDALAETSEFARKWVPFVKYDYIEPRAPFY
FFKIRINALSVSKALKVPEEGWIMODGTPWPFGN

```

Db 191 -----VRP-----EPEQFAKDCGAL----- 205
QY 496 RDHPGMIQVFLHSGGLDTGEGNELPRLVVYSREKRPFGQHHKKAGAMNALVRVSAVLNG 555
Db 206 -----YIGRVDS-----HAKAGNLNHAIK-----RTSG 229
QY 556 QYMLNDCDHYINNSKAVREAMCFMLDPNLGPQVCYVQFPQDFGIDRNDRYANRNTVFF 615
Db 230 DYILILDCDH-IPTRAFLOIAMGMWVADR---KIALMQTPHHFYSPD----- 272
QY 616 DINLRGLDGOQPVYVGTGCFVNRNTAIYGYEPPKAKKPGFLASLCGKKKASKKRSS 675
Db 273 -----PFQNLAVGYRTP----- 285
QY 676 DKRKNKHVDSSVPVFNLEIDIEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFAVSTLME 735
Db 286 -----PEGNL---FYGVIOQDND-----FWDATFFCGSCAI- 313
QY 736 YGVPQSSTPESLLKEAIIHVISGCEYEDKSEWTEIGWIYGVSTEDILTGFKMHARGWRSV 795
Db 314 -----LREAAIESIG-----GFVETVTEDAHTALRMQRGWSTA 348
QY 796 YCMPKRPAPKGSAPINLSRNLQVLRWALGSEVILFSRHCPWYGGRLKFLERFAYIN 855
Db 349 YL--RIPVASGLATERLTHIGQMRWARGMIOI--FRVDNPM---LGGGLKGLQRLCYLS 402
QY 856 TTYPLTSLPLLYCYLPAICLTGKFIPEISNLASIFIALFSIFATGILEMRS-- 913
Db 403 AMTSEFFAIPRVIFLASPLAFPFQGNII-----AASPLAVLAYIAPHMFHSTA 451
QY 914 -GVGIDSWRNEQFW-VIGGISAHLEAVFQGLKVLGIDTNTFTVTSKA---NDEGDPA 968
Db 452 TAAKYNKGR-YSEWSEVETTWALFLVRVTIITLMPKSGKNVTEKGVLEEFEDLG 510
QY 969 ELYMFKWTTLLIPTTILINMVGWVAGTSYAINSGYQSGWPLFG--KLFFAFWVIVHLV 1026
Db 511 ATY-----PNIIF-----AGIMTLGLLIGLFEITFHF----- 537
QY 1027 PFLKGLMGRONRPTIVVWA-----VILLASI 1053
Db 538 ---NQLAGIAKRAYLNCIWMASLILLAAI 566

RESULT 7
US-08-960-048-10
; Sequence 10, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960, 048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029, 987
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 756
; TYPE: PRT
; ORGANISM: acetobacter xylinum
US-08-960-048-10
Query Match 4.1%; Score 237.5; DB 4; Length 756;
Best Local Similarity 20.1%; Pred. No. 3 5e-13;
Matches 156; Conservative 89; Mismatches 232; Indels 301; Gaps 31;
QY 280 VLRVLSIFLHYLRNLNVR-NAYPLWLLSVI---CEIWFALSWILDQFPKWFPIINRETY 335
Db 80 VLSALVSLRYLTWRLTETLDNTWIGILGVILLMAELYALMYFLSYFQTIQPLHRAP- 138

QY 336 LDRLALRYDREGEPSQAAVDIFVSTVDPLKEPPIVITANTVLSILAVDVPVKVCSYSD 395
Db 139 ---LPLPDNDWDT---VDIFPTVD---EQLSIVLRVLGALGIDWPPDKVNVYILD 188
QY 396 DGASMLTFDALAETSEFAKRWVFVKYDIEPRAPFEYFCOKIDYLDKQVQPSFVKDRA 455
Db 189 DG-----VRP----- 193
QY 456 MKREYEERFIRINALVSKALVPDEGWIMQDGTWPQGNNTDRDHPGMIQVFLHSGGLDTE 515
Db 194 ---EPEQFAKDCGAL-----YIGRVDD 213
QY 516 GNEPLRVLYVSREKRPFGQHHKKAGAMNALVRVSAVLNGQYMLNDCDHYINNSKAVRE 575
Db 214 ---SAHAKAGNLNHAIK-----RTSGDYILILDCDH-IPTRAFLOI 250
QY 576 AMCFMLDPNLGPQVCYVQFPQDFGIDRNDRYANRNTVFFDINLRGLDGOQPVYVGTGC 635
Db 251 AMGMWADR---KIALMQTPHHFYSPD----- 274
QY 636 VFNRTAIYGYEPPKAKKPGFLASLCGKKKASKKRSDKKKNKHVDSSVPVFNLE 695
Db 275 PFQNLAVGYRTP-----PEGNL-- 292
QY 696 IEEGVEGAGPDEKSVLMSQMSLEKRFQGSAAFAVSTLMEYGGVPOSSPESLLKEAIIH 755
Db 293 -FYGVIOQDND-----FWDATFFCGSCAI- 323
QY 756 ISCYEDKSEWTEIGWIYGVSTEDILTGFKMHARGWRSVYCMKRPAPKGSAPINLSR 815
Db 324 IG-----GFVETVTEDAHTALRMQRGWSTAYL--RIPVASGLATERLTH 368
QY 816 LNQVLRWALGSEVILFSRHCPWYGGRLKFLERFAYINTTYPLTSLPLLYCYLPAI 875
Db 369 IGQMRWARGMIOI--FRVDNPM---LGRGLKGLQRLCYLSAMTSEFFAIPRVIFLASPLA 434
QY 876 CLLGKGTIMPEISNLASIFIALFSIFATGILEMRS---GVGIDSWRNEQFW-VIGG 931
Db 425 FLFAGQNI-----AAPLAVAYALPHMFHSIATAAKVNKGR-YSEWSEVYE 472
QY 932 ISAHLEAVFQGLKVLGIDTNTFTVTSKA---NDEGDFAELY-MFKWTTLLIPTPTI-- 985
Db 473 TTMALFLVRVTITVTLFPKSGKNVTEKGVLEEFEDLGATYPTIIFATIMMGGLLIGL 532
QY 986 --LII--NMVGWVAGTSYAINSGYQSGWPLFGKLFFAFWVIVHLVPELKL-MGRONR 1038
Db 533 FELIVRFNQLDVIARNAYLLNCA-----WALISLIILFAAIVGRETK 575

RESULT 8
US-09-147-236-3
; Sequence 3, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-0PCT
; CURRENT APPLICATION NUMBER: US/09/147, 236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Acetobacter xylinum
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or


```

QY 754 HVISCGVEDKSEWGTETGNIWYSVTEDIULTGFKMHARGWRSVYCMKPKRPAKGSAPINLS 813
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 320 EQIG-----GFATQVTTVEDAHTALKMQLGWSTAYL--RIPLAGGLATERLI 364
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 814 DRLNQVLRWALGSVEILFSRHCPLWTVGYGGRKLFLEFAYINTTTPUTSLPLVYICLP 873
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 365 LHTQQRVWARGMLQI-FRIDNPL---FGRGLSWGQRCLYLSAMTSFLFAVPRVIFLSSP 420
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 874 AICLLGKFTIMPEISNLASWIFALFSLPFGATGILEMRWSGVG-----IDENWRNEQFW-V 928
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 LAFLFQGNII-----AASPLALLAYAIPIHM-FHAVGTASKINKGWR-YSEWSE 467
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 929 IGGISAHFLAVFOGLKVLVLAGIDTNTFTVSKANDEGDFAEILYMKFWTTLIPPTTILII 988
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 468 VYETTMALFLVRVITYTLLSPSRGKFNVTDKGLLEKGYFDLG-----AVYENIILGL 520
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 989 NMVGWVAGTSAINSQYSGWGLFGKLFPAFWIVHLYPFLKGL-MGRQ-----NRTPT 1040
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 521 IMFGLARGVYELSFGLDQIAERAYLLNSAWAMLSLIILAAIAVAGRETQOKRNSHIP 580
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1041 TIVIV 1045
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 581 ATIPV 585
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-08-119-773-4
; Sequence 4, Application US/08119773
; Patent No. 5460942
; GENERAL INFORMATION:
; APPLICANT: Chou, Janice Y.
; APPLICANT: Lei, Ke-Jian
; APPLICANT: Shelly, Leslie L.
; TITLE OF INVENTION: GLUCOSE-6-PHOSPHATASE: THE GENE AND
; TITLE OF INVENTION: PROTEIN AND RELATED MUTATIONS
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,773
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-175
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..357
; OTHER INFORMATION: /label= R to C at 83

```

OTHER INFORMATION: /note= "The G-6-Pase amino acid sequence mutated from Arg to Cys at position 83."
OTHER INFORMATION: from Arg to Cys at position 83.

Query Match 2.0%; Score 114; DB 1; Length 357;
Best Local Similarity 19.1%; Pred. No. 0.043;
Matches 72; Conservative 59; Mismatches 137; Indels 108; Gaps 16;

QY 751 EAHVISGDEKSWGTEIGWISGVTEDILTGFKMHARGWRSVYCMKPRPAKGSAPI 810
DB 14 QSTHYLVNQYQSDQW----FILSVIADLRNAPYVLPFW---FHLQEAVGIKLLWA 65
QY 811 NLSDRNLQVLRWALGSVELLSRHCHPLWYGGGRUKFLERFAYINTTYPYTSPLP--- 867
DB 66 VIGDWLNLVFKW-----ILFQ-CPYWM-----VLDTDYISNTSVPLIKQF 105
QY 868 -----VYCIU-PAICLLTGKFMPEIS----- 888
DB 106 PVTCTGPGSPSGHAMGTAGYVYVNVMTSLISFOGK-IKPYRFLCNVILWLFQWVQL 164
QY 889 ---NLASVFIALFLSIFATGILEMRWSGVGIDEMWRN-----EOWVWIGGISAHLF 937
DB 165 NVCLSRIVIAAHFPHQVAVGLV---SGIAVTETSHISYINASLKKYFL---ITFPLF 217
QY 938 AVFQGLLVLAGIDNFTVTSKANDDEGDEAFELYMFKWTTLLIPPTILLINMGVAVGT 997
DB 218 SFAIGFYLLKGLGVLDLTLWLEKAGRWCEQP-----WVHDTTPFASLLKLN-GLTFLG 271
QY 998 SVAINSG-----YOSWGLFLGKFFAFWVHLVYFPLKGLMGRNRTTIVVWAV 1048
DB 272 GIALNSMYRESCKGLSKWLPFLRSSIVASLVLLHVFDSLK---PPSQVELVFYVLSF 327
QY 1049 LLASIFLLWVRVDPF 1064
DB 328 CKSAVPLASVSVIPY 343

RESULT 12

US-08-669-656A-8
Sequence 8, Application US/08669656A
Patent No. 6451554

GENERAL INFORMATION:

APPLICANT: Wood, John N.
APPLICANT: Akopian, Armen N.
TITLE OF INVENTION: Ion Channel
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZENECA Pharmaceuticals
STREET: 1800 Concord Pike, P.O. Box 15437
CITY: Wilmington
STATE: Delaware
COUNTRY: USA
ZIP: 19850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/669,656A

FILING DATE: 24-JUN-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Hohenschutz, Liza D.

REGISTRATION NUMBER: 33,712

REFERENCE/DOCKET NUMBER: PHM.70086

TELEPHONE: (302) 886-7466

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1957 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein
us-08-669-656A-8

Query Match 1.9%;
Best Local Similarity 17.5%;
Matches 207; Conservative 107; Mismatches 100; Indels 57; Gaps 57;

QY 66 TKYRHRGSPAIRGEEDG--- 120
DB 439 TSLQSHSGSPKLANERRE 490
QY 121 -----SCNVCHPKYDSCE 162
DB 491 LSSGRRRASHGSHVFFHFRAP 550
QY 163 GASPDHMHMSPTGNISRRAP 220
DB 551 -----RSPLE 571
QY 221 IPMTNGTSIAPSEGRATDIDA 279
DB 572 VPTGELTAGAP-EGPAL- 609
QY 280 VLRVWLISIFLHYRLTNPNR 338
DB 610 AMSVSI-----MTSVIEE 655
QY 339 LALRYDREGSPSQAADVIFV 395
DB 656 MAL-----FE 691
QY 396 DGASMLTFDALAE----- 736
DB 692 -----AFDAMLQAGNIVFTV 766
QY 446 QPSFVKDRAMKREYEEFKIR 505
DB 737 VTVSLLESASAKG----- 779
QY 506 LGHSGGLDTEGNELPRLVYVS 560
DB 780 IGNS----- 810
QY 561 LD----- 579
DB 811 EDYCKRKDGVSVMNGEKLWHE 870
QY 580 LMDPNLGPQVCYVQFPQRFQ 628
DB 871 LTVNVLGNLVNLNLF----- 912
QY 629 -----VYVGC 961
DB 913 ALARIQVLGHRASRAIASYIS 972
QY 662 GKKKASKSKKRSDDKKSKNK 709
DB 973 NLTPALSSPKENHGDFTDF 1032
QY 710 -----SVL 741
DB 1033 EQLPQVKCNHQAARSPASM 1092
QY 742 SSTPEILLKEAIVH----- 781
DB 1093 CPDPEELIRKIPDELADDER 1144
QY 782 LTGFKMHARGWRSVYCMKPR 836
DB 1145 KTCYRIVEHSHWFSFIIFMIL 1204
QY 837 L-WYGYGGRUKFLERFAYINT 880

Db 1205 LKVVAVGKKYFTNACWLDLVLNLSLTLAKILEYSDVASIKALRTLRALPLRLALS 1264
QY 881 KF-----IMPEISNLASIWFTA-LFLSIFATGILEMRWSGV-----GID 918
Db 1265 REFEGMRVVVDALVGAIPSMNVLLVCLIFPWLIFSMGVNLFAGKFSKCVTRNNPFSNVN 1324
QY 919 EWRNEQ-----FW-----IGGISAHFLFAVFOGLLKLVL-AGIDTNTF 955
Db 1325 STVWNKSECHNQNSTGTFWVNVKVFNDVAMGYLALLOVATFKGMDIMYAAVDG-E 1383
QY 956 VTSKANDEGDFALYMFKWTLLIPPTTILINM-VGVV 994
Db 1384 INSOPNWN-----NLYMYLFVVVFIIFGGFTLNLFGVI 1419

RESULT 13

US-08-669-656A-2
; Sequence 2, Application US/08669656A
; Patent No. 6451554
; GENERAL INFORMATION:

; APPLICANT: Wood, John N.
; APPLICANT: Akoplan, Armen N.
; TITLE OF INVENTION: Ion Channel
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals
; STREET: 1800 Concord Pike, P.O. Box 15437
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19850

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,656A
; FILING DATE: 24-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: PHM.70086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-7466
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1957 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-669-656A-2

Query Match 1.9%; Score 111; DB 4; Length 1957;
Best Local Similarity 18.2%; Pred. No. 1.5;
Matches 212; Conservative 152; Mismatches 380; Indels 422; Gaps 59;

QY 66 TKYKRHRGSPAIRGEED-----DTDADDGSDPNYPASCTEDQKQKIAIDRMRSWRMTGG 120
Db 439 TSLQSHSGSLASKANERPRVKRSVSEGS-----TDDNRSPQSDPYNQRMSFLG 490
QY 121 -----SGNYGHPKYDSGEI-----GLSKYD-----SGEIPRGVPSVTSNMQSGEIP 162
Db 491 LSSGRRRASHGVSFHFRAPOSQISFPDGTDPDGVFPHGQESRRGSILLGRGAGQTGPLP 550
QY 163 GASPDHMSPTGNISRAPFPVNVHSPNPREF--SGSIGNVANKERVYDGHKMKQDKGA 220
Db 551 -----RSLPL--QSPNPGRRHGEQGLG----- 571
QY 221 IPMTNGTSIAPSEGRAATOIDASTEYNMEDALINDETQPLSKVPFIASSKIN-PYRMVI 279
Db 572 VPTGELTAGAP-EGPAL-----HTGKSKFSLSAGYLNEPFRAGR 609

QY 280 VLRLVVLISFLHRLNPNVRNAYPLWLLSVICEIWFALSUIL-DQFPKWFPIINRETYLDR 338
Db 610 AMSVSI-----MTSVIEELEESKLCPCCLISFAQKYLWECCKPKRKF-----K 655
QY 339 LALRYDREGPSOLAADVIFSVTDPLKBPPI---TANTVLSILAVDYPVKVSCYVSD 395
Db 656 MAL-----FELVTDFAELTITLCIVVNTVFAME-HYPMTD-----691
QY 396 DGASMLTFDALAE-----TSEFARKWVPFVKKYDIEPRAPPEFYFCQKIDYLDKVK 445
Db 692 -----AFDAMLOAGNIVTFVFTMEMAFKIIAF-----DPYIFYQKKWNFDQCVI 736
QY 446 QPSFVKDRRAMKREYEEFKIRINALYSKALKVPEEGWIMQDGTWPNGNTRDHPGMIOVF 505
Db 737 VTVSLLELSASKG-----SLSVLRTLRLLR---VFKLAKSWFTLNT-----LIKI- 779
QY 506 LGHSGGLDTGEGNELPRLVYVSREKRFQFHKKAGAMNALVRVSAVLT-----NQGYMLN 560
Db 780 IGNS-----VGAIGNLTFILAIIVFIFALVGKQLLS 810
QY 561 LD-----CDHYINN-----SKAVREAMC---F 579
Db 811 EDYGCRRKDGYSVWNGEKLWHMCDFFHSLVWFRILCGEWIENMWYCMVEYSQKSLILF 870
QY 580 LMDPNLGPQVCYVQFPQDFGIDRNDRYANRNTVFFD---INLR-GLDGIQ-----626
Db 871 LTVWVIGNLVVL-----NLFIALLLNSFSADNLTAPEDDGEVNNLQALARIQVLGHRASR 926
QY 627 -GPVYVGTGCVFNRTAI---YGYEPPKAKK-----PGFLASLGGKKKASKSKRRSS 675
Db 927 ASASYISSHCRFHWPKVETQLGKMPKPLTSEAKNHATDAVSAAGVNLTKPALSSPKENH 986
QY 676 DKKSNNKHVDSSYPV---FNLEDIEGVGAG-----PDDEK-----709
Db 987 GDFITDPNVWVSPVIAEGESDLDEEDMEQASQSSQWQEDPKGQOQELPQVOKCENHQA 1046
QY 710 ----SVLMSQMSLEKRFQGS-----AAEFAVSTLMEYGGVQPSSTPESLLK---BA 752
Db 1047 ARSPASMMSSDDLAPYLGESKRRKSDQVPAEGVDDTSSSEGSTVDCPDPEELRKIPEL 1106
QY 753 IHVI-----SCGYE-----DKSEWGEIHWIYGSVTEIDLTGFKKHARGWSRV 795
Db 1107 AHDLEDPDCDFREGCTRRCPCCNVNTSKSPWAT--GW-----QVRKTCYRIVEHSWFES 1158
QY 796 YCMKRPAPKGSAP-----INLSDRLNQVLRWALGSVEILFSRHCLP-WVGYGRULKLE 849
Db 1159 FIIFWILLSSGALAFEDNLYEKPVRKSVLEYTDRTVTFIFVEMLLKWAIVGFKKYFTN 1218
QY 850 RFAYINTTI--YPLTSL--PLLVCILPATCLL-----TGKF-----882
Db 1219 AWCWLDLIVNISLTLIAKILEYSDVASIKALRTLRALPLRLALSFRFEGMRVVVDALVG 1278
QY 883 IMPISINLASIWFTA-LFLSIFATGILEMRWSGV-----GIDWNRNEQ-----925
Db 1279 AIPSIMVLLVCLIFLWIFLIFSMGVNLFAGKFSKCVTRNNPFSNVNNTVWNKSECHNQ 1338
QY 926 ----FW-----IGGISAHFLFAVFOGLLKLVL-AGIDTNTVTSKANDEGDFAE 969
Db 1339 STGHFFWVNVKVFNDVAMGYLALLOVATFKGMDIMYAAVDG-EINSOPNWN-----N 1393
QY 970 LYMKFWTLLIPPTTILINM-VGVV 994
Db 1394 LYMYLFVVVFIIFGGFTLNLFGVI 1419

RESULT 14
US-08-485-568A-6
; Sequence 6, Application US/08485568A
; Patent No. 5589382
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Narva, Kenneth E.


```
;
; FILING DATE: 21-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/918,345
; FILING DATE: 21-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/558,738
; FILING DATE: 27-JUL-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,698
; FILING DATE: 16-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,403
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/999,053
; FILING DATE: 14-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA48DD2.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-590-554A-6

Query Match      1.9%; Score 108; DB 2; Length 1167;
Best Local Similarity 19.7%; Pred. No. 1.2;
Matches 139; Conservative 77; Mismatches 214; Indels 276; Gaps 36;

QY 411 EFARKW-----VPFK--KYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRAMKREYEE 462
Db 250 QFAERKWSVNYDESINOTKVDLQRIOD--YSTVTSTTFEKFETLNPSSKESVKNYNR 307
QY 463 FKIRINALYSKALKV-----PPEGWIMODGT-----PWPGNNTRDHPGM 501
Db 308 Y---VRSMTLQSLDIAATWPTLDNVPNSVDIQLDQRLVFSVDVAGPWEGNDNIT-SNI 363
QY 502 IQVFLGHSGGLD-TEGNEPLRLVYVSREKRPGFQHHKAGAMNALRVYSAVLTNGQYMLN 560
Db 364 IDVLTPINTGIGFQESSDLRKFY-PRIELOSQMFH-----GQYV-- 402
QY 561 LDCDHYNNKAVREAMCFLMDPLNGPQVYVQFPQRFQDGRDNDRYANR----- 610
Db 403 -----NSKSVEH-----CY-----SDGLKLN--YANKTITAGVSNID 432
QY 611 -----NTVPFDINRLGLDGIQGPVYVG-----TGCVFNRATAYG 644
Db 433 ESNQNNKHNGYPVINSPTIDINVS-----QNSOYLDLNSVMVNGQKVTCG----- 479
QY 645 YEPPIKAKPGFLASLCGGKKKSKKSSDKKKSNKHDV-----SSVPVFNLED 695
Db 480 --SPLSSNGNSNAALPNQKINVIYS-VOSNDKPE--KHADTYRKWGYMSSHIP-YDL-- 531
QY 696 IEEGVEGAGFDDKE--SVLMSQMSLEKRFQSOAAFV-----ASTLMYEGGVQPSSTP 745
Db 532 VPENVIGDIDPTKQPSLLKLGKFPKQYDSIAIYVSEPLGANAVKLTSTQVLMQEVTN 591
QY 746 ESKLKEAIVH--ISCGYEDKSEWGEIGWYGSVTE-----ILTFKMHARGWRSYIC 797
Db 592 QTTQKYRIRIRVATGDTAASIWFHILIGPSGNDLTNEGHNFSSVSRNKMVFQGNNGKY- 650
```

```
QY 798 MPKRPAKGSAPINLSDRNLNQVLRWALGSVEILFSRHCPWLWYGYGRLKFLERFAYINTT 857
Db 651 -----VNLILTDSIELPSSGQQTILIONTSQDLFLDRIEFIS-- 687
QY 858 IYPLTSLPLLVYCILPAICLLTGKFMPE-----ISNLASIWFIALFLSFATGILEMRWS 913
Db 688 -LPSTSTP-----TSTNFVEPESLEKIINQVN---QLFSSSSQTEL----- 724
QY 914 GVGIDEMWRNEQWPWVIGGISAHL---FAVFOGLLKVLAGIDNTFTVTSKA----- 960
Db 725 -----AHTVSDYKIDQVVLKYNALSDDVFGVEKKALKRLVNAQAKQ 764
QY 961 -----NDEEGD-----FAELYMFKWTTLLIPPTTI 985
Db 765 LSKARNVLVGGNFEKGHEWALSREATMVANHELFGDHLPPPTL 810
```

Search completed: February 19, 2003, 16:37:43
Job time : 47 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2003, 16:36:16 : Search time 31 Seconds
(without alignments)
890.090 Million cell updates/sec

Title: US-09-900-237-30
 Perfect score: 5778
 Sequence: 1 MDGDADALKSGRHAGDVQC.....VDPFTTRLGPNIOCTGINC 1080

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:	Minimum	Match	0%
	Maximum	Match	100%
	Listing	first	45 s

```

Database : Published_Applications_Aa:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	5778	100.0	1080	10	US-09-900-237-30	Sequence 30, Appl
2	4507	78.0	1065	10	US-09-900-237-33	Sequence 33, Appl
3	4455.5	77.1	881	10	US-09-838-539-8	Sequence 8, Appl
4	4021.5	69.6	1091	10	US-09-900-237-26	Sequence 26, Appl
5	3844.5	66.5	1165	10	US-09-900-237-8	Sequence 8, Appl
6	3820.5	66.1	1039	10	US-09-900-237-14	Sequence 14, Appl
7	3818	66.1	1086	10	US-09-900-237-10	Sequence 10, Appl
8	3818	66.1	1148	10	US-09-900-237-4	Sequence 4, Appl
9	3477	60.2	974	10	US-09-838-539-6	Sequence 6, Appl
10	3404.5	58.9	701	10	US-09-900-237-32	Sequence 32, Appl
11	3234	56.0	793	10	US-09-900-237-18	Sequence 18, Appl
12	3038.5	52.6	740	10	US-09-900-237-24	Sequence 24, Appl
13	2816	48.7	685	10	US-09-900-237-31	Sequence 31, Appl
14	2811	48.7	685	10	US-09-838-539-7	Sequence 7, Appl
15	2689	46.5	506	10	US-09-900-237-20	Sequence 20, Appl
16	2627.5	45.5	610	10	US-09-900-237-16	Sequence 16, Appl
17	2044.5	35.4	431	10	US-09-900-237-28	Sequence 28, Appl
18	1418.5	24.6	320	10	US-09-900-237-6	Sequence 6, Appl
19	1379.5	23.9	304	10	US-09-900-237-2	Sequence 2, Appl

ALIGNMENTS

```

RESULT 1
US-09-900-237-30
; Sequence 30, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/9
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,84
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 30
; LENGTH: 1080
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-900-237-30

```

	Query Match	100.0%;	Score 5778;	DB 10;	Length 1080;	
	Best Local Similarity	100.0%;	Pred. No. 0;			
	Matches 1080;	Conservative	0;	Mismatches	0;	Indels
						Gaps
Qy	1	MDGDADALKSRHGAGDVCQICADGLGTLTDGDVFTACDVCGRFFVPCRYEHERKEGTQA	60			
Db	1	MDGDADALKSRHGAGDVCQICADGLGTLTDGDVFTACDVCGRFFVPCRYEHERKEGTQA	60			
Qy	61	CLOCKTKYKRRHGSPAIRGEGEDTDADDGSDFNYPASGTEDQKQTADNRKSRNMVTGG	120			
Db	61	CLOCKTKYKRRHGSPAIRGEGEDTDADDGSDFNYPASGTEDQKQTADNRKSRNMVTGG	120			
Qy	121	SGNVGHPKYDGSGETGLSKYDSGEIPRGVVPVSVTNQSMQSGEIPGASPDHMMSPGTNISR	180			
Db	121	SGNVGHPKYDGSGETGLSKYDSGEIPRGVVPVSVTNQSMQSGEIPGASPDHMMSPGTNISR	180			
Qy	181	APFVYVNHSPNPREFGSGISGNVAKKRVDCGWKKQDKGALPMPFNGTISIAPSEGRAADTI	240			

Db 181 APPYYNHSNPSREFSGSIGNVAMKERVGDGKMKDKGAIPMTNGTSTAPSEGRATDI 240
QY 241 DASTEXMEDALLNDETROPLSRKVPFIASSKINPYRMWIVLRVLVLSIFLHYRLTNPNRN 300
Db 241 DASTEXMEDALLNDETROPLSRKVPFIASSKINPYRMWIVLRVLVLSIFLHYRLTNPNRN 300
QY 301 AYPLWLLSVCEIWFALSILWIDQPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFS 360
Db 301 AYPLWLLSVCEIWFALSILWIDQPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFS 360
QY 361 TVDPLKEPIVANTVLSILAVDYPVDKVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
Db 361 TVDPLKEPIVANTVLSILAVDYPVDKVCYVSDDGASMLTFDALAETSEFARKWVPFV 420
QY 421 KKYDIEPRAPEFYFCOKIDYLDKQVOPSFVKDRRAKREYEFKIRINALVSKALVPEE 480
Db 421 KKYDIEPRAPEFYFCOKIDYLDKQVOPSFVKDRRAKREYEFKIRINALVSKALVPEE 480
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPRLVYVVSREKRPFGQHHKKAG 540
Db 481 GWIMQDGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPRLVYVVSREKRPFGQHHKKAG 540
QY 541 ANNALVRVSAVLNGQYMLNLDCHYINNSKAVREACMFMDPNLPQVCYVQFPQREFG 600
Db 541 ANNALVRVSAVLNGQYMLNLDCHYINNSKAVREACMFMDPNLPQVCYVQFPQREFG 600
QY 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
Db 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYEPPKAKKPGFLASL 660
QY 661 CGGKKKASKSKRSDDKSKNSKHDVSSVPFNLEDEEGVEGAGFDDKSKVMSQMSLEK 720
Db 661 CGGKKKASKSKRSDDKSKNSKHDVSSVPFNLEDEEGVEGAGFDDKSKVMSQMSLEK 720
QY 721 RFQSQSAFVASTLMEYGGVQPSQSTPESLKEAHLVISCYEDKSENGTEIGWYGSVTE 780
Db 721 RFQSQSAFVASTLMEYGGVQPSQSTPESLKEAHLVISCYEDKSENGTEIGWYGSVTE 780
QY 781 ILTGFKMHARGWSVYCMKPRPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLWYG 840
Db 781 ILTGFKMHARGWSVYCMKPRPAFKGSAPINLSDRNLQVLRWALGSVEILFSRHCPLWYG 840
QY 841 YGRLKFLERFAYINTIYPLTSLPLLVYCIPLPAICLLTGKFTMPERISNLASIFIALFL 900
Db 841 YGRLKFLERFAYINTIYPLTSLPLLVYCIPLPAICLLTGKFTMPERISNLASIFIALFL 900
QY 901 SIFATGILEMRWSGVGIDENWNEQFWIIGGISAHLPAFVQGLLKVLAGIDTFTVTSKA 960
Db 901 SIFATGILEMRWSGVGIDENWNEQFWIIGGISAHLPAFVQGLLKVLAGIDTFTVTSKA 960
QY 961 NDEEGDFAELYMEKWTLLIPPTTILINMGVYVAGTSYAINSGYOSWGLPFGKLPFAFW 1020
Db 961 NDEEGDFAELYMEKWTLLIPPTTILINMGVYVAGTSYAINSGYOSWGLPFGKLPFAFW 1020
QY 1021 VIVHLYPFLKGLMGRONRPTIIVWAVLLASIFSLWLRVDPFTTFLAGPNQTQCGINC 1080
Db 1021 VIVHLYPFLKGLMGRONRPTIIVWAVLLASIFSLWLRVDPFTTFLAGPNQTQCGINC 1080

RESULT 2

US-09-900-237-33
; Sequence 33, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13

; PRIOR APPLICATION NUMBER: 09/
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-900-237-33

Query Match 78.0%;
Best Local Similarity 78.4%;
Matches 836; Conservative

QY 19 CQICADGLGTLTGDGVFTACDV
Db 20 CQICSDNKGTVTDGDFVACDI
QY 79 GEEGDGTDADGGS-DFNYPASG
Db 80 GKDEGDGLADEGTVFNYP--
QY 138 KYDSGEIPRGYVPSVTNSQ-MS
Db 128 -----EVSHNHLPRLTSRQDTS
QY 196 FSGSIGNVANKERVGDGKMKQD
Db 183 DPVGLGNVANKERVGDGKMKQD
QY 256 ETRQPLSRKVPFIASSKINPYRM
Db 239 EAQPLSRKVSIFSSRINPYRM
QY 316 ALSWILDQPKWPIINRETYLD
Db 299 ALSWILDQPKWPIINRETYLD
QY 376 VLSILAVDYPVDKVCYVSDDG
Db 359 VLSILAVDYPVDKVCYVSDDG
QY 436 QKIDYLDKVKQPSFVKDRRAK
Db 419 AKIDYLDKVKQPSFVKDRRAK
QY 496 RDPHGMIOVFLGHSGLDTEGN
Db 479 GDHPGMIQVFLGHSGLDTEGN
QY 556 QYMLNLDCHYINNSKALREAK
Db 539 PFILNLDCHYINNSKALREAK
QY 616 DINLRGLDGIQGPVYVGTGCVF
Db 599 DINLRGLDGIQGPVYVGTGCVF
QY 674 SSDKSKNSKHVDSSVPFNLEDE
Db 659 ESKDKKSGRHTDSTVPVFNLD
QY 734 MEYGGVQPSQSTPESLKEAHLV
Db 719 MENGVPVPSATPENLLKEAHLV
QY 794 SVTCMKRPAFKGSAPINLSDR
Db 779 SIYCMKPLPAFKGSAPINLSDR
QY 854 INTTIYPLTSLPLLVYCIPLPAI
Db 839 VNTTIYPTITSIPLLMYCTLLAV

3
; re 4507; DB 10; Length 1065;
; id. No. 0;
; Mismatches 107; Indels 26; Gaps 8;
CRPCYBHERKEGTQACLOCKTKTKRHRGSPAIR 78
CRPCYBHERKEGTQACLOCKTKTKRHRGSPAIR 79
QKIDRMRSMRMTNGSGNVGHPKYDSEIGLS 137
EKISERMGLGHLTRGKGEEMGEPOYDK----- 127
ASPDHMMSTGNISRRAPFPY-VNHSNPSRE 195
ASPERLSVSTIAGGRKRLPYSDVNQSPNRIV 182
MTNGTSTAPSEGRATDIIDASTEYNMEDALLND 255
V----STQAASE-RGGVDIDASTDILADEALLND 238
LVVLSIFLHYRLTNPNVAYPLWLLSVCEIWF 315
LVILCLFLHYRITNPVPNAPALVSVCEIWF 298
YDREGPSQLAAVDIFVSTVDPLKEPIVANT 375
YDREGPSQLAAVDIFVSTVDPLKEPIVANT 358
FDALAESEFARKWVPVKKYDIEPRAPERYFC 435
FESLAETSEFARKWVPFCKYSTEPRAPERYFA 418
FKIRINALVSKALKVPEEGWIMQDGTWPNGNT 495
FKIRINALVSKALKVPEEGWIMQDGTWPNGNT 478
VYVSREKRPFGQHHKKAGANNALVRVSAVLNG 555
VYVSREKRPFGQHHKKAGANNALVRVSAVLNG 538
PNLGPVYVQFPQRFQIDRNDRYANRNTVFF 615
PNLGPVYVQFPQRFQIDRNDRYANRNTVFF 598
YGYEPPIKA--KKPGFLASICGGKKKASKSKKR 673
YGYEPPIKVHKHKKPSLLSKLCGSSRKNKSKAK 658
EGAGFDEKSKVMSQMSLEKRFQSAFVASTL 733
EGAGFDEKSKVMSQMSLEKRFQSAFVASTL 718
EDKSENGTEIGWYGSVTEIDLTGFKMHARGWR 793
EDKSDMGMEIGWYGSVTEIDLTGFKMHARGWR 778
RWALGSVEILFSRHCPLWYIGRGLKFLERFAY 853
RWALGSVEILFSRHCPLWYIGRGLKFLERFAY 838
KFTMPERISNLASIFIALFUSIFATGILEMRWS 913
QFIIPQISNLASIFUSLFLSIFATGILEMRWS 898

```
QY 914 GVGIDWNRNEQFWIGISAHLEFAVFOGLKVLKVLADIDNTFTVTSKANDSEGDFAEIYMF 973
Db 899 GVGIDWNRNEQFWIGISAHLEFAVFOGLKVLKVLADIDNTFTVTSKANDSEGDFAEIYLF 958
QY 974 KWTLLIPPTTLLIINMVGWVAGTSYAINSGYQSWGPLFGKLFGLFAFWVIVHLVYFPLKGLM 1033
Db 959 KWTLLIPPTTLLIINMVGWVAGTSYAINSGYQSWGPLFGKLFGLFAFWVIVHLVYFPLKGLM 1018
QY 1034 GRQNRPTIIVWAVLLASIFSLWVRVDPPTTRLAGPNTQTCGNC 1080
Db 1019 GRQNRPTIIVWVSVLLASIFSLWVRIDPFTSRVTPGDIILECGINC 1065

RESULT 3
US-09-838-539-8
; Sequence 8, Application US/09838539
; Patent No. US20020129401A1
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/03/US
; CURRENT APPLICATION NUMBER: US/09/838,539
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/029,987
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 881
; TYPE: PRT
; ORGANISM: Oryzae sativa
US-09-838-539-8

Query Match 77.1%; Score 4455.5; DB 10; Length 881;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;

QY 201 GNVAKERVDMKQKGAIPMTNGTSIAPSEGAATPDIDASTYNNMEDALLNDETRQP 260
Db 1 GNVAKERVDMKQKGAIPMTNGTSIAPSEGAATPDIDASTYNNMEDALLNDETRQP 60
QY 261 LSRKVPASSKINPYRMVITLRLVLSIFLHYLTNPVRNAYPLWLLSVCEIWFALSWI 320
Db 61 LSRKVPASSKINPYRMVITLRLVLSIFLHYLTNPVRNAYPLWLLSVCEIWFALSWI 120
QY 321 LQDFPKWFFINRETYLDRALRYDREGPSQLAAVDFVSTVDPLKEPPIVANTVLSIL 380
Db 121 LQDFPKWFFINRETYLDRALRYDREGPSQLAAVDFVSTVDPLKEPPIVANTVLSIL 180
QY 381 AVDYPVDKVCYSDDGASMLTFDALAETSEFARKWVPFKYDIEPRAPEEFCOKIDY 440
Db 181 AVDYPVDKVCYSDDGASMLTFDALAETSEFARKWVPFKYDIEPRAPEEFSQIDY 240
QY 441 LKDKVQPSVKORRAMKREYEEFKIRINALYSKALKVPPEEGWIMQDGTWPGNNTRDPG 500
Db 241 LKDKVQPSVKORRAMKREYEEFKIRINALYSKALKVPPEEGWIMQDGTWPGNNTRDPG 300
QY 501 MIQVFLHSGGLDTGEGNELPRLVYVSREKRPFGQHHKAGAMNALVRVSAVLNNGQYMLN 560
Db 301 MIQVFLHSGGLDTGEGNELPRLVYVSREKRPFGQHHKAGAMNALVRVSAVLNNGQYMLN 360
QY 561 LDCDHYINNSKAVREAMCFMIDPDLGPOVCYVQFPQRPQDGDIDRNDYRANRNTVFEDINLR 620
Db 361 LDCDHYINNSKALREAMCFMIDPDLGPOVCYVQFPQRPQDGDIDRNDYRANRNTVFEDINLR 420
QY 621 GLDGIQGPVYVGTGVFNRTALYGYEPPIKAKKPG-FLASLCGKKKASKSKSSDKKK 679
Db 421 GLDGIQGPVYVGTGVFNRTALYGYEPPIKAKKPG-FLASLCGKKKASKSKSSDKKK 480
```

```
QY 680 SNKHVDSSVPYFNLEIDEEGVGAGFDDEKSVLMSQMSLEKRFQSQSAFVASTIMEYGGV 739
Db 481 SNKHVDSSVPYFNLEIDEEGVGAGFDDEKSVLMSQMSLEKRFQSQSAFVASTIMEYGGV 540
QY 740 PQSSTPESLLKEAIVHISCGYEDKSEWGTGTEIGWYIGSVTEDITLTGPKMHARGWRSYVCM 799
Db 541 PQSSTPESLLKEAIVHISCGYEDKTEWGTGTEIGWYIGSVTEDITLTGPKMHARGWRSYVCM 600
QY 800 KRPAPKSGAPINLSDRLNQLRWALGSEVILFSRHCHPLWYGYGGRGLKFLERFAYINTIY 859
Db 601 KRPAPKSGAPINLSDRLNQLRWALGSEVILFSRHCHPLWYGYGGRGLKFLERFAYINTIY 660
QY 860 PLTSLPLLVYCILPAICLLTGKFIIMPETISNLASIFLALFSLFATGILEMRWSGVGIDE 919
Db 661 PLTSLPLLVYCVLPAICLLTGKFIIPISNPFASIFLALFSLFATGILEMRWSGVGIDE 720
QY 920 WWRNEQFWVIGGISAHLEFAVFOGLKVLKVLADIDNTFTVTSKANDSEGDFAEIYFKWTTLL 979
Db 721 WWRNEQFWVIGGISAHLEFAVFOGLKVLKVLADIDNTFTVTSKANDSEGDFAEIYFKWTTLL 780
QY 980 IPPTTILIIINMVGWVAGTSYAINSGYQSWGPLFGKLFGLFAFWVIVHLVYFPLKGLMGRQNR 1039
Db 781 IPPTTILIIINMVGWVAGTSYAINSGYQSWGPLFGKLFGLFAFWVIVHLVYFPLKGLMGRQNR 840
QY 1040 PTIVIVWAVLLASIFSLWVRVDPPTTRLAGPNTQTCGNC 1080
Db 841 PTIVIVWAVLLASIFSLWVRIDPFTSRVTPGDTQTCGNC 881

RESULT 4
US-09-900-237-26
; Sequence 26, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 26
; LENGTH: 1091
; TYPE: PRT
; ORGANISM: Impatiens balsamia
US-09-900-237-26

Query Match 69.6%; Score 4021.5; DB 10; Length 1091;
Best Local Similarity 68.1%; Pred. No. 0;
Matches 749; Conservative 138; Mismatches 160; Indels 53; Gaps 20;

QY 2 DGDADALSKSGRHAGDVQICADGLTDLGDVFTACDVCRFPVCRPCYERHERKEGTQAC 61
Db 24 DSDGQPKPLKEANGQICQICGDTVGKTSATGDTFVACNECGFPVCRPCYEVERKDGQCC 83
QY 62 LOCKTKYKRRHSGPAIRGEEDDTDADD-GSDFNVPASCTEDQKQKTADRRSHRMNTGG 120
Db 84 POCKTKYKRRHSGPAIRGEEDDTDADD-GSDFNVPASCTEDQKQKTADRRSHRMNTGG 135
QY 121 SCNVGHKPYDSGEISLKYVDSGEIPRGVYPSVTNS-QMSGEIPGASPDHMM-----SPTG 175
Db 136 ----GDQDDIELSVSSSRHDESQRP---VPLLTGHSHSVSGEIP--TFDNHSIRTSPIG 186
QY 176 NTSRRAPFPYVN-HSP-----NPSREF-SGSIGNVANKERVDMKQKGAIPMTNG 226
Db 187 PVEK--SIPYIDPQPAVRVIRIIVDPKSLNSYGLGNVDKERVGWKLKQEKNNVQMTS- 243
```

QY 227 TSAPSEGRATDIDASTYNMEDALLNDETROPLSRKVPFIASSKINPYRMVIVLRLVVL 286
Db 244 ---RYPEGKGTBTGSGNEELQMAA--DDIRQPMRSIVPISSTHLTPRVVILRLIIL 298
QY 287 SIFLHYRLINPVNAPYLLWSVICELWALSIILOQFKWFPINREYTLDRALRYDRE 346
Db 299 GFLQRCRTHPVKDAVPLWLTSCVCEWFWALSIILOQFKWFPINREYTLDRALRYDRE 358
QY 347 GEPSQLAAVDIFVSTVDPLKEPPIVANTVLSILAVDYPDKVSCVSDGASMLTFDAL 406
Db 359 GEPSQLAPIDVSVSTVDPLKEPPIVANTVLSILAVDYPDKVSCVSDGASMLTFDAL 418
QY 407 AETSEFARKWVPVKKYDIEPRAPRYFCOKIDYLDKQVQSFVDRRAKREYEFKIRI 466
Db 419 SETAEFAKKWAPCKKHSDEPRAPRYFCOKIDYLDKQVQSFVDRRAKREYEFKIRI 478
QY 467 INALVSKALKVPEEGWIMODGTPWPGNNTRDHPGMIQVFLHSGGLDTEGNEPLRLVYVS 526
Db 479 INALVAKAKVPEEGWIMODGTPWPGNNSRDHPGMIQVFLHSGGLDTEGNEPLRLVYVS 538
QY 527 REKPGFQHHKAGAMNALVRVSAVLNTNGQYMLNDCDHYINNSKAVREAMCFMDPNLG 586
Db 539 REKPGFQHHKAGAMNALVRVSAVLNTNGAYLLNVDCHYFNNSCKLKEAMCFMDPNLG 598
QY 587 PQVCYVQFQRFQDIDRNDRYANRNTVFFDINLRLDGIQGVYVGTGCVFNRTAIYGE 646
Db 599 KKTICYVQFQRFQDIDRNDRYANRNTVFFDINLRLDGIQGVYVGTGCVFNRTAIYGE 658
QY 647 PPKAK--KPG--FLASLCGGKKKASKKRSDDKSKNKHVDSVVPVFNLEIDIEGVEGA 703
Db 659 PVLTEDELFNLIILKSCCSRSKGGKGNKKYIDKNRALKRTSTAPIFNMEDIEGIE-- 716
QY 704 GFDEKSVLMSQMSLEKRFQSAAFVASTLMEYGGVQPSSTPESLLKEAHHVISCYEDK 763
Db 717 GYDERSERLMAQ--SYEKRFQSPVLIAATFMEQGLPPSTNSATLLKEAHHVISCYEDK 775
QY 764 SEWGTGIEGIVSVTEEDITLTKFMHARGWRSVYCMKPAFKGSAPIINLSRLNQVLKWA 823
Db 776 TENGKEIGWIVSVTEEDITLTKFMHTRGWISYCMPPPAFKGSAPIINLSRLNQVLKWA 835
QY 824 LGSVELFRHRPLWTVGYGGRKLFLERFAYINTIYPTLSLPLLVYCIPLAICLLTGKFI 883
Db 836 LGSIEILLRHRCPWTVGYGGRKLFLERLAYINTIYPTLSLPLLVYCIPLAICLLTGKFI 895
QY 884 MPEISNLASTWTFALFLSFATGILEMRWSGVIDEWNRNEQFWIGGSAHLFAVFOGL 943
Db 896 VPEISNASTWTFALFLSFATGILEMRWSGVIDEWNRNEQFWIGGSAHLFAVFOGL 955
QY 944 LKVLADIDNTFTSKANDEEGDFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINS 1003
Db 956 LKVLADIDNTFTSKANDEEGDFAELYMFKWTTLLIPPTTILINMVGVVAGTSYAINS 1015
QY 1004 GYOSWGPFLGKLEFAFWVILVHLPFLKGLMGQRNTPTIIVIVAVALLASIFSLWVRVDP 1063
Db 1016 GYOSWGPFLGRLFFAFLWVILVHLPFLKGLMGQRNTPTIIVIVAVALLASIFSLWVRIDP 1075
QY 1064 FT---TRLAGPNIQTGGINC 1080
Db 1076 FTSDSTKARG---QCGIDC 1091

RESULT 5

US-09-900-237-8
; Sequence 8, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14

; PRIOR APPLICATION NUMBER: PCT/US98/015871

; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/083,333
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 1165
; TYPE: PRT
; ORGANISM: Zea mays
US-09-900-237-8

Query Match

66.58%

Best Local Similarity 65.4%

Matches 723; Conservative

Mismatches 175; Indels 69; Gaps 15;

QY 9 KSGRHGADV-----CQICAE 62

Db 96 ESGRAGGGAARRAPCOICG 155

QY 63 QCKYKXKRRGSPAIRGEGDE 121

Db 156 QCRTRYKRLKCPRVAGDEEE 215

QY 122 GNUGHPKYDSCIEIGLSKYDSG 175

Db 216 GD-AHP-----CG 257

QY 176 NISRAPPPYVN-----HST 227

Db 258 GSKRIHLPFPADPNLPVQPRSS 309

QY 228 SIAPSEGRAADTIDASTEYNN 287

Db 310 QHVRSEGGDWDGDA----- 363

QY 288 IFLYRLTNVNRNAYPLWLLS 347

Db 364 FFHYRYMHPAKDAFALWLIS 423

QY 348 EPSQLAAVDIFVSVVDPLKEP 407

Db 424 QPSQLAPIDFVSVVDPLKEP 483

QY 408 ETSFARKWVPVKKYDIEPR 467

Db 484 ETSFARKWVPVKKYDIEPR 543

QY 468 NALYSKALKVPEEGWIMODGT 527

Db 544 NALYAKAKVPEEGWIMODGT 603

QY 528 EKRGFQHHKAGAMNALVRV 587

Db 604 EKRCYNHKKAGAMNALVRV 663

QY 588 QVCVQFPQRFQDIDRNDRYA 647

Db 664 KVCYVQFPQRFQDIDRNDRYA 723

QY 648 PIKAKP-----GFLA 695

Db 724 P-KTKKPPSRTCNCPKWCLS 782

QY 696 IEEGVGAGFDDKSVLMSQ 755

Db 783 IDEGAPGA--DIEKAGIVNO 840

QY 756 ISCYEDKSEWGTGIEGIVGS 815

Db 841 ISCYEDKDWGKEIGWYGS 900

QY 816 LNOVLWALGSVEILFSRHC 875

US-09-900-237-10

Query Match	66.1%	Score 3818;	DB 10;	Length 1086;			
Best Local Similarity	64.74;	Pred. No. 1.8e-315;					
Matches 716;	Conservative 141;	Mismatches 175;	Indels	Gaps 17;			
Qy 2	DGDADALKSRG	HAGDVQV	ICADGLGTTLDG	VFTACDVC	RFPVCRPCY	EHKEGTQAC	61
Db 23	DGDGP	-KPPRQNG	QVQICGDDV	GLAPGGD	PFVACNECAF	VCVRCDCY	81
Qy 62	LQCKTKYK	KHRGSPAIR	EGEGDDTD	DADDGSD	FNYPASGT	EDQOKTAD	121
Db 82	POCKTRYK	KLKGCQR	VTGDEEDG	VDDLDNEF	NDWG-----	HDQSVAESML	128
Qy 122	GNYPHKPY	DGSGEIGLSK	YDSSGIPRGY	-----	VPSVTNSQMS	GEIPGASD	174
Db 129	--YCHMSY	CRGG-----	DPNGAPQAF	QNLNPVLL	TNGQVMD	DIP---PEQ	177
Qy 175	-GNISRRAP	PPVYN-----	HSPNPSREF	SG-SIGNVAM	KERV	DGKKQDKG	225
Db 178	GGGKRIH	PLADPSLP	VQPSRMDP	SKDLAAY	GVSGVAM	KERNWKQ	235
Qy 226	GTSIAPSE	GRAATDIDA	STEYNMED	ALLNDET	ROP	LGRKVP	285
Db 236	-----	NDGGDDG	DDA-----	DLP	LMDEAR	QOLSRK	282
Qy 286	LSITFLYR	TNPVRNAY	PLWLISV	ICEWAL	SWILOQ	PKWPPIN	345
Db 283	LGFFHYR	VYWHVP	NDAPAL	WLISV	ICEWAL	SWILOQ	342
Qy 346	EGEPSQ	LAADVIF	VSTVDP	LKEPPI	VTANTV	LSILA	405
Db 343	EGQPSOL	APTDFV	VSTVDP	LKEPPL	VTNTV	LSILS	402
Qy 406	LATSEFAR	KWPFV	VKKYDIE	PAPEFF	YFCKID	PLDKV	465
Db 403	LSETSEF	AKWPF	VCKRYNIE	PAPEW	YFQKID	PLDKV	462
Qy 466	RINALVSK	ALVPEEG	IMODGT	PWPGN	TRDHPG	MIQV	525
Db 463	RINALVAK	QVPEEG	WTMODGT	PWPGN	VRNDRH	FCMIQV	522
Qy 526	SREKRP	GFOHHK	KAGAMN	ALVRV	SAVLN	NGOYML	585
Db 523	SREKRP	GNHKK	KAGAMN	ALVRV	SAVLN	APYLLN	582
Qy 586	GPQVCY	VQPPQ	RDGD	INDR	YANRTV	FFDIN	645
Db 583	GKVCYV	QPPQ	RDGD	IDR	YANRV	FFDIN	642
Qy 646	EPPIKAKP	-----	GFLASL	CGKKK	AKSKRS	DDKK--	693
Db 643	DAP-KYK	PPSR	CNCWPK	WCECC	CFGNR	KQKTK	701
Qy 694	EDIEEG	VEGAF	DEKSV	ILMSOM	SELEK	RFGSOA	753
Db 702	GEIDEA	APGA--	-ENEKAG	VYNQOK	LEKX	FGOSS	759
Qy 754	HVISC	GVDEK	SEMTG	EIGWY	SVTE	DIITG	813
Db 760	HVISC	GVDEK	TDGKE	IHWY	SVTE	DIITG	819
Qy 814	DRLNQV	RNALG	SVELF	SRHC	PLWY	GGRLK	873
Db 820	DRLNQV	RNALG	SVELF	SRHC	PLWY	GGRLK	879
Qy 874	AICLLT	GKFT	IMP	SIN	SLAS	IFAL	933
Db 880	AICLLT	GKFT	IMP	SIN	SLAS	IFAL	939
Qy 934	AHLFV	EQGLL	KVLG	AD	TNFTV	SKAN	993
Db 940	SHLFV	EQGLL	KVLG	AD	TNFTV	SKAN	998

QY 539 AGAMALYRVSAVLNTNGOYMLNLDGDHYINNSKAVRAMCFMNDPNLGPQVCYVQPPORF 598
DB 596 AGAMALYRVSAVLNTNADYLNLDGDHYINNSKAKRAMCFMNDPLLGGKVCYVQPPORF 655
QY 599 DQIDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGVFNRTAIYGYEPPKAKKP----- 654
DB 656 DQIDRNDYANRNVVFFDINMKGLDGIQGPVYVGTGVFNRTAIYGYEPPKAKKP----- 714
QY 655 -----GFLASLGGKKKASKSKRSDK-----KSKNHVDSSVVPVFNLEDIEGVEGAG 704
DB 715 CNCFKWCICCCCFNRRKTKTKTSKPKFKIKKLFKKENQAPAYALGEIDEAAPGA- 773
QY 705 FODERSVLMSOMSLKREKFGQSAAFVASTLMEYGGVPOSSPESLLKEAIIHIVISGVEDKS 764
DB 774 -ENERASIVNOOKLEKFGQSSVFVASTLLENGGTLKASASPASLLKEAIIHIVISGVEDKT 832
QY 765 ENGTEIGNYGVTEDITGFKMHARGWRSVCMKRPAPKGSAPINLSRLNOVLKAL 824
DB 833 GWKDGIGYGVSTEDITGFKMHARGWRSVCMKRPAPKGSAPINLSRLNOVLKAL 892
QY 825 GSVEILSRHCPWYGGRLKFLERFAYINTIYPLTSLPLLYCYCILPAICLLTGKFI 884
DB 893 GSIEFFSNHCPWYGGRLKFLERFAYINTIYPLTSLPLLYCYCILPAICLLTGKFI 952
QY 885 PEISNLASTWLTALPSTAFGILEMRWSGVGIDENRNEQFVWIGISAHLFAVQGLL 944
DB 953 PELNVASLWFSLEFICFATSIEMRWSGVGIDENRNEQFVWIGISAHLFAVQGLL 1012
QY 945 KYLAGIDTNTVTSKANDEEDGFAELYMFKWTTLLIPTTLLIINMGVVGAGSYAINSG 1004
DB 1013 KYIAGVDTSFTVTSKGGDE-EFSELYTFKWTLLIPTTLLIINMGVVGAGSYAINSG 1071
QY 1005 YOSWGLPKGKFFAFWVTHLYPLKGLMGRQNRTPITIVVAVLLASIFSLMWVRDPF 1064
DB 1072 YESWGLPKGKFFAFWVTHLYPLKGLMGRQNRTPITIVVAVLLASIFSLMWVRDPF 1131
QY 1065 TTRLAGPNQTCGINC 1080
DB 1132 LAKDDGPLEEGCLDC 1147

RESULT 9
US-09-838-539-6
; Sequence 6, Application US/09838539
; Patent No. US20020129401A1
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/03/US
; CURRENT APPLICATION NUMBER: US/09/838,539
; PRIOR FILING DATE: 2001-04-18
; PRIOR FILING DATE: 1996-10-29
; PRIOR APPLICATION NUMBER: 08/960,048
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-838-539-6

Query Match 60.2%; Score 3477; DB 10; Length 974;
Best Local Similarity 61.5%; Pred. No. 1.4e-286;
Matches 657; Conserved 129; Mismatches 176; Indels 106; Gaps 15;

QY 78 RGEEDDTDADDGSDFNYPASGTEQKQKIADRMRSRMTNGSGSNVG-HPKYDSGEIGL 136
DB 59 -----DENLDD-----VERATGQSTMAA-----HLNKSDQVGIHARHS-----SV 96
QY 137 SKYDSGEIPRGYVSVTNQSOMSGEIPGASPDHMHMSPTGNISRRAPFPYVNHSPNPREF 196
DB 97 STLDS-----EM 103
QY 197 SSGISGNVAKERVDCWKKQDKGAIPMTNGTSIAPSGRAATDIDASTEYNMEDALLNDE 256
DB 104 AEDNGNSLWKNRVSWEKKNKKKPPAT-----KVERAEIPEOQMEDKPADPA 154
QY 257 TRQPLSRKVPASSKINPVMVILVLSIFLHYRLTNPNRYNAYPLWLISVCEIWEFA 316
DB 155 S-QPLSTIIPKSLAPYRTVIIMRLIILGLFFHYRTNPVDSAFGLWLTSVCEIWEFA 213
QY 317 LSWILDQPPKFPINREYIIDLRLRYDREGEPSQLAAVIFVSTVDPLKEPPIVANTV 376
DB 214 FSWILDQPPKFPVNRRETYIDRLSARYEREGEDELAAVDFVSTVDPLKEPPLITANTV 273
QY 377 LSLIADYVPDKVSCYVSDDGASMLTFDALAETSEFARKWVPFVKYKVDIEPRAPEFYFCQ 436
DB 274 LSLIALDYVPDKVSCYISDDGAMLTPELVETADAFARKWVPFCKKFSIEPRAPEFYFCQ 333
QY 437 KIDYLUKVKVQSFVKDRAMREYEEKIRINALVSKALKVPBEGWTIMQDGTWPQGNTR 496
DB 334 KIDYLUKVKVQSFVKERRAMRDYEEKIRINALVAKAQTDPDEGWTIMQDGTWPQGNTR 393
QY 497 DHPGMIOVFLGHSGGLDTEGNEPLRVYVSRKRPQGHKKKAGAMNALVRVSAVLNTNG 556
DB 394 DHPGMIOVFLGSGARDIEGNEPLRVYVSRKRPQGHKKKAGAMNALVRVSAVLNTNAP 453
QY 557 YMLNDDCHYINNSKAVREACMFLMDPNLGPQVCYVQPPORFQIDRNDYANRNTVFFD 616
DB 454 FTLNDDCHYINNSKAVREACMFLMDPNLGPQVCYVQPPORFQIDRNDYANRNTVFFD 513
QY 617 INLRGLDGIQGPVYVGTGVFNRTAIYGYEPPKAKKPGLASLCGGKSKSKRSD 676
DB 514 YNMKGIDGIQGPVYVGTGVFNRTAIYGYEPPKAKKPGLASLCGGKSKSKRSD 573
QY 677 KKKSNKHVDSSVVPVFNLEDIEGVEGAGFDD-EKSVLMSOMSLKREKFGQSAFVASTLME 735
DB 574 LYRDAKREELDAAIFENLEID-----NYDEYERSMLISQTSFEKTFGLSVFESTIME 627
QY 736 YGGVPOSSTPESLLKEAIIHIVISCGVEDKSEWTEIGWYGVSTEDILTGFKMHARGWRSV 795
DB 628 NGVAESANPSTLKEAIIHIVISCGVEKTAGKEIGWYGVSTEDILTGFKMHARGWRSV 687
QY 796 YCMKRRPAFKGSAPINLSRLNOVLKGLMGRQNRTPITIVVAVLLASIFSLMWVRDPF 854
DB 688 YCMKRRPAFKGSAPINLSRLNOVLKGLMGRQNRTPITIVVAVLLASIFSLMWVRDPF 747
QY 855 NTIYPLTSLPLLYCYCILPAICLLTGKFIPEISNLASIFLAFATGILEMRWSG 914
DB 748 NTIYVPTSLPLIACVSLPAICLLTGKFIPTLNLASVFLGLFSLIYVAVLEMRWSG 807
QY 915 VGIDENRNEQFVWIGISAHLFAVQGLLKVLAGIDTNTVTSKANDEGDFAEALYMFK 974
DB 808 VSIEDLWRNEQFVWIGISAHLFAVQGLLKVLAGIDTNTVTSKANDEGDFAEALYMFK 866
QY 975 WTTLLIPTTLLIINMGVVGAGSYAINSGVPOSSPESLLKEAIIHIVISGVEDKS 1034
DB 867 WTTLLIPTTLLIINMGVVGAGSYAINSGVPOSSPESLLKEAIIHIVISGVEDKS 926
QY 1035 RQNRTPITIVVAVLLASIFSLMWVRDPFTTTRLAGPNI-QTC-GINC 1080
DB 927 RQNRTPITIVVAVLLASIFSLMWVRDPFTTTRLAGPNI-QTC-GINC 974

RESULT 10
US-09-900-237-32
; Sequence 32, Application US/09900237
; Patent No. US20020120124A1

```

; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-900-237-32

Query Match      58.98; Score 3404.5; DB 10; Length 701;
Best Local Similarity 88.28; Pred. No. 1.2e-280;
Matches 618; Conservative 56; Mismatches 24; Indels 3; Gaps 2;

QY 383 DYPVDKVCYVSDGASMLTSEFARKWVPFVKYKDYDIEPRAPFEYFCQKIDYK 442
DB 1 DYPVEKVCYVSDGASMLTSEFARKWVPFVKYKDYDIEPRAPFEYFCQKIDYK 60
QY 443 DKVQSFVKDRAMKREYEFEKIRINALVSKALVPEGWIMQDGTWPNGNTRDHPGMI 502
DB 61 DKVQSFVKERRAMKREYEFEKIRINALVSKALVPEGWIMQDGTWPNGNTRDHPGMI 120
QY 503 QVFLGSGGLDTEGNEPLRVVSRKRPQHHKAGAMNALVRVSAVLNGQVMLNLD 562
DB 121 QVFLGSGGLDTEGNEPLRVVSRKRPQHHKAGAMNALVRVSAVLNGQVMLNLD 180
QY 563 CDHYINNSKAVREAMCFMNDPLNGQVYVQFPQRFQDIDRNDRYANRNTVFEDINLRGL 622
DB 181 CDHYINNSKALREAMCFMNDPLNGQVYVQFPQRFQDIDRNDRYANRNTVFEDINLRGL 240
QY 623 DGIOGPVYVGTGCVFNRTAIGYEPPIKAK--KPGFLASLCGG-KKASKSKKSSDKKK 679
DB 241 DGIOGPVYVGTGCVFNRTAIGYEPPIKAK--KPGFLASLCGG-KKASKSKKSSDKKK 300
QY 680 SNKHVDSPVPFNLEDIEEGVEGAGFDDKESVLMQSOMLEKRFQGSAAFAVASTLMYEGV 739
DB 301 SGKHVDSPVPFNLEDIEEGVEGAGFDDKESVLMQSOMLEKRFQGSAAFAVASTLMYEGV 360
QY 740 QOSTPPEILLKEAHHVISCYEDKSEWGTETGWIYGSVTEDLITGFKMHARGWRSYICMP 799
DB 361 QOSTPPEILLKEAHHVISCYEDKSEWGTETGWIYGSVTEDLITGFKMHARGWRSYICMP 420
QY 800 KRPAFGSAPINLSRLNQVRLWALGSVEILFSRHCPLWYGSYGGRLKFLERFAYINTIY 859
DB 421 KRPAFGSAPINLSRLNQVRLWALGSVEILFSRHCPLWYGSYGGRLKFLERFAYINTIY 480
QY 860 PLTSLPLLVCYLPAICLLTKGFMPEISNLSASTWFIALFLSIPATGILEMRWSGVGIDE 919
DB 481 PVTATPLMLCYLPAICLLTKGFMPEISNLSASTWFIALFLSIPATGILEMRWSGVGIDE 540
QY 920 WRNRQFVWVIGGISAHLEFAVFGQLKVLGAGIDTNTFTVTSKANDDEGDFAEILMFKWTLL 979
DB 541 WRNRQFVWVIGGISAHLEFAVFGQLKVLGAGIDTNTFTVTSKANDDEGDFAEILMFKWTLL 600
QY 980 IPTTILLINMVGVSAGTSYAINSGYQSWGDFLFGKLFPAFVVIHLVFLKGLMGRQNR 1039
DB 601 IPTTILLINMVGVSAGTSYAINSGYQSWGDFLFGKLFPAFVVIHLVFLKGLMGRQNR 660
QY 1040 PTIVLVAVLLASIFSLVLRVDPRTTRLAGPNITQTCGINC 1080
DB 661 PTIVLVAVLLASIFSLVLRVDPRTTRLAGPNITQTCGINC 701
```

```

RESULT 11
US-09-900-237-18
; Sequence 18, Application US/09/900,237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: 60/092,844
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Glycine max
US-09-900-237-18

Query Match      56.08; Score 3234; DB 10; Length 793;
Best Local Similarity 73.38; Pred. No. 4.4e-266;
Matches 581; Conservative 87; Mismatches 87; Indels 14; Gaps 5;

QY 296 NPVRNAYPLMLLSVCEIWFAN 355
DB 4 HPVNDAYGLMLTSVCEIWFAN 63
QY 356 DIFVSTVDPLKEPPPIVNTANTV 415
DB 64 DVFVSTVDPMKEPLIANTV 123
QY 416 WVPVKYKDYDIEPRAPFEYFCQ 475
DB 124 WVPCKYKDYDIEPRAPFEYFCQ 183
QY 476 KVPEEGIMQDGTWPNGNTR 535
DB 184 KVPEDGWTMQDGTWPNGNTR 243
QY 536 HKKAGAMNALVRVSAVITNGQ 595
DB 244 HKKAGAMNALVRSAIITNAP 303
QY 596 QRFQDIDRNDRYANRNTVFED 655
DB 304 QRFQDIDRNDRYANRNTVFED 363
QY 656 -----FLASLCGGKKKAS 707
DB 364 KTCNCWPKWCCLCCGSRKKKY 419
QY 708 EKSVMQSOMLEKRFQGSAAFA 767
DB 420 EKSVMQSOMLEKRFQGSAAFA 479
QY 768 TEIGWIYGSVTEDLITGFKMH 827
DB 480 KEVGIYGSVTEDLITGFKMH 539
QY 828 EILFSRHCPLWYGGRLKFL 887
DB 540 EILFSRHCPLWYGGRLKFL 599
QY 888 SNLASIMFIALFLSIFATGIL 947
DB 600 SNLASIMFIALFLSIFATGIL 659
QY 948 AGIDTNTFTVTSKANDDEGDF 1007
```

```
Db 660 AGVNTFTVTSKAAD-DGEFSELYIFKWTSLIPMTLLIMNIVGVVGISDAINNGYDS 718
QY 1008 WGPLFGKLPFAFWVIVHLYPFLKGLMGRNRPPTIVIVWAVLLASIFSLLLWVRVDPFTTR 1067
Db 719 WGPLFGKLPFAFWVIVHLYPFLKGLMGRNRPPTIVIVWAVLLASIFSLLLWVRVDPFTTR 778
QY 1068 LAGPNITQTCGINC 1080
Db 779 -DGPVLEICGLNC 790

RESULT 12
US-09-900-237-24
; Sequence 24, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Picramnia pentandra
US-09-900-237-24

Query Match 52.6%; Score 3038.5; DB 10; Length 740;
Best Local Similarity 74.1%; Pred. No. 1.6e-249;
Matches 552; Conservative 97; Mismatches 79; Indels 17; Gaps 6;

QY 345 REGEPSOLAVIDFVSTVDPLKEPPIVTANTVLSILAVDPYDKVSCYVSDGASMLTFD 404
Db 3 KEKPSLACLDIFVSTVDPMKEPPITANTVLSILAVDPYDKVTCYVSDGASMLTFE 62
QY 405 ALAETSEFARKWPFVKVYDIEPRAPEFYFCOKIDYLDKVKOPSFVKDRRAMKREYEFK 464
Db 63 ALSETSEFARKWPFCKKFSIEPAPENWFSQMDYLNKNVHPSFVRERRAMKREYEFK 122
QY 465 IRINALVSKALKVPEEGWIMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVY 524
Db 123 VRINGLVAMAQKVPEDGWTMQDGTWPNGNVRDHPGMIQVFLGHNGVRDVEGNELPRLTY 182
QY 525 VSREKRPGFQHHKAGAMNALVRSAVLNNGOYMLNDCDHYINNSKAVREACFLMDPN 584
Db 183 VSREKRPGFQHHKAGAMNSLRVSAVSNAPYILNVDCDHYINNSKALREACFLMDPT 242
QY 585 LGPOVCYVQPPQRFQDGRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIVG 644
Db 243 SGKKLCYVQPPQRFQDGRNDRYANRNVFFDINMKGLDGIQGPVYVGTGCVFRRVLYG 302
QY 645 YEPIKAKKPG-----FLSLCGGKK-KASKSKRSDDKKSKNHVDSSVPFNLED 695
Db 303 YDAPVTTKSPGKACNCPWKLCCCGGSRKSKKSKKPKK-----EKKSKNREASQKHALEN 358
QY 696 IEEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFAVSTLMYEGVGPQSGSTPESILKEAIVH 755
Db 359 IEEGM--GGLNSEKSCETTPLEKKEFGQSPFVASTLLEDGGVPQDAPALLKEAIOY 416
QY 756 ISCYEDKSEWGEIGNIYGSVTEIDILTGFKMHARGWRSYVCMKRPAPFKGSAPINLSDR 815
Db 417 ISCYEDKTEWGEVGIYGSVTEIDILTGFKMHCHGWRSYVCMKRPAPFKGSAPINLSDR 476
QY 816 LNOVLRWALGSVELFSRHCPLWYGVGGRLLKFLERFAYINTTIYPLTSLPLLYVCTLPAL 875
```

```
Db 477 LHOVLRWALGSVELFSRHCPLWYGVGGLKWLERSYVSSVYVWPTSLPLLYVCTLPAL 536
QY 876 CLITGKREIMEISNLSINFIATGILEMWSGVGIDEMWRNEQFVWIGGISAH 935
Db 537 CLITGKREIVEISNYASILFMLFIATATSIEMOMGGVGDMMWRNEQFVWIGVSSH 596
QY 936 LEAFVQGLLKVLAGIDTNTFTVTSKANDGEDGFAELYMFKWTLLIPPTTILINIWMGVYA 995
Db 597 LEALFQGLLKVLAGVNTFTVTSKAAD-BGDFSELYLFKWTLLIPPTTLLINIIVGVV 655
QY 996 GTSYAINSGYQSGWGLPGLKLFPAFWIVHLYPFLKGLMGRNRPPTIVIVWAVLLASIPS 1055
Db 656 GVSDAINNGYDSWGLPFLGRFFAFWIVHLYPFLKGLLQKQDRTPTIIVWSILLASILT 715
QY 1056 LLWVRVDPFTTRLAGPNITQTCGINC 1080
Db 716 LLWVRINPFVSR-DGPVLEVCGLNC 739

RESULT 13
US-09-900-237-31
; Sequence 31, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-900-237-31

Query Match 48.7%; Score 2816; DB 10; Length 685;
Best Local Similarity 74.0%; Pred. No. 1.1e-230;
Matches 509; Conservative 78; Mismatches 77; Indels 24; Gaps 6;

QY 414 RKWVPVFKYKVDIEPRAPEFYFCOKIDYLDKVKOPSFVKDRRAMKREYEFKIRINALVSK 473
Db 1 RRWVPFCKKHNVPEPRAPEFYFENEKIDYLDKVKHPSFVKERRAMKREYEFKVRINALVAK 60
QY 474 ALKYPEEGWIMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSRKRPGF 533
Db 61 AQKPESGWVMDGTWPNGNTRDHPGMIQVILGSAGALVDGKELPRLVYVSRKRPGF 120
QY 534 QHKKKAGAMNALVRSAVLNNGOYMLNDCDHYINNSKAVREACFLMDPNLGPQVCYVQ 593
Db 121 QHKKKAGAMNALVRSAVLNAPFLNLDCHYINNSKAMREACFLMDPQFGKLCYVQ 180
QY 594 FPQRFQDGRNDRYANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIVGYEPPKAKK 653
Db 181 FPQRFQDGRNDRYANRNVFFDINMLGLDGIQGPVYVGTGCVFNROALYGPDPVSEKR 240
QY 654 -----PGFLSLCGGKKKASKSKRS-----SDKKK---SNKHVDSVVPFNL 693
Db 241 PKMTCDCWPSWCCCCCGSRKSKKKGKGLLGLLYGKKKMMGKNYVYKGSAPVFDL 300
QY 694 EDIEGVEGAGFDD-EKSVLMSQMSLEKRFQGSAAFAVSTLMYEGVGPQSGSTPESILKEA 752
Db 301 BEIEEGLE--GYEELEKSTLMSQNFKEKRFQSPVFIASTLMENGGLPEGTNSTSLIKEA 358
QY 753 IHVISCYEDKSEWGEIGNIYGSVTEIDILTGFKMHARGWRSYVCMKRPAPFKGSAPINL 812
```

Db 359 IHVISCYEEKTEWKEIGWYGSVTEIDLTGFKMHCGRWKSVCVCKRPAFKGSAPINL 418
Qy 813 SDRNLQVLRWALGSVEILFSRHCPLWYGYGRIKFLERFAINTIYPLTSLPLLYVCIL 872
Db 419 SDRHLQVLRWALGSVEILFSRHCPLWYGYGRIKFLERFAINTIYPLTSLPLLYVCIL 478
Qy 873 PAICLLTGKFTMPISNLASIFWALFISFATGILEMRWGSVGVDEWRNEQFWIGGI 932
Db 479 PAVCLLTGKFIPLSNLTSWFLALFSLIATGVLELRWGSVGVDEWRNEQFWIGGI 538
Qy 933 SAHLFAVFOGLLKVLAGIDTNTVTSKANDDEGDAELYFVKWTLLIPPTTILINMV 992
Db 539 SAHLFAVFOGLLKVLAGIDTNTVTSKANDDEGDAELYFVKWTLLIPPTTILINMV 597
Qy 993 VVAGTSYAINSGYSGWPLFGKLEFAFWVILHLYPFLKGLMGRNQRTPTIIVLWVLLAS 1052
Db 598 VVAGTSYAINSGYSGWPLFGKLEFAFWVILHLYPFLKGLMGRNQRTPTIIVLWVLLAS 657
Qy 1053 IFSLWVRVDPFTTTLRAGPNIQTGCINC 1080
Db 658 IFSLWVRVDPFTTTLRAGPNIQTGCINC 685

RESULT 14
US-09-838-539-7
; Sequence 7, Application US/09838539
; Patent No. US20020129401A1
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/03/US
; CURRENT APPLICATION NUMBER: US/09/838,539
; CURRENT FILING DATE: 2001-04-18
; PRIOR FILING DATE: 1996-10-29
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-838-539-7

Query Match 48.7% Score 2811; DB 10; Length 685;
Best Local Similarity 74.1%; Pred. No. 2.9e-230; Indels 24; Gaps 6;
Matches 509; Conservative 76; Mismatches 76;

Qy 413 ARKWWPFVKKYDIEPRAPEFYFCQKIDYLDKQVPSFYKDRRAMKREYEEFKIRINALVS 472
Db 1 ARRWVPFCCKHNVPRAPPEFYNEKIDYLDKQVPSFYKDRRAMKREYEEFKIRINALVA 60
Qy 473 KALKVPEGLWMDGTWPGNTRDHPGMIOVFLHSGDLTTEGNEPLRVYVREKRP 532
Db 61 KAQKKPEGWYMDGTWPGNTRDHPGMIOVFLHSGDLTTEGNEPLRVYVREKRP 120
Qy 533 FQHHKAGAMNALVRVSAVLNTNGOYMLNDCDHYINNSKAVREACMFLMDPQGVCIY 592
Db 121 YQHHKAGAMNALVRVSAVLNTNAPFLNDCDHYINNSKAVREACMFLMDPQGVCIY 180
Qy 593 QPQRFQDIDRDRYANRNTVFFDINLRLDGIQGVVYVGTGCVFNRTAIYGYEPPKAK 652
Db 181 QPQRFQDIDRDRYANRNTVFFDINLRLDGIQGVVYVGTGCVFNRTAIYGYEPPKAK 240
Qy 653 K-----PGFLASGCGKKKSKRS-----SDKKK---SNKHVDSSVPVFN 692
Db 241 RPKWTCDCPWSWCCCCGGRKKKKKKKKGLGLGLGLGLGLGLGLGLGLGLGLGLGLGL 300
Qy 693 LEDIEEGVEGAGFDD-BKSVLMSOMSLKFRGQSAFVASTLMEYGGVQPSSTPESLLKE 751
Db 181 LEDIEEGVEGAGFDD-BKSVLMSOMSLKFRGQSAFVASTLMEYGGVQPSSTPESLLKE 751

Db 301 LEEIEGLE--GYEELEKSTLM
Qy 752 AIHVISGYEDKSWGTEIGWII
Db 359 AIHVISGYEKEKTEWKEIGWII
Qy 812 LSDLNQLVLRWALGSVEILFSR
Db 419 LSDLNQLVLRWALGSVEILFSR
Qy 872 LPAICLLTGKFIPEISNLASII
Db 479 IPAVCLLTGKFIPTLSNLTSTV
Qy 932 ISAHLEFAVFOGLLKVLAGIDTNI
Db 539 VSAHLFAVFOGLLKVLAGIDTNI
Qy 992 GVAGTSYAINSGYSGWPLFG
Db 598 GVAGTSYAINSGYSGWPLFG
Qy 1052 SIFSLWVRVDPFTTTLRAGPNI
Db 658 SIFSLWVRVDPFTTTLRAGPNI

RESULT 15
US-09-900-237-20
; Sequence 20, Application US/09
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cell
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: 0
; CURRENT FILING DATE: 2001-07
; PRIOR FILING DATE: 1998-07-14
; PRIOR FILING DATE: 1999-07-13
; PRIOR FILING DATE: 1999-07-13
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (88)
US-09-900-237-20

Query Match 46.5%
Best Local Similarity 99.6%
Matches 504; Conservative

Qy 575 EAMCFMDPNLGPQVCYVQFPQ
Db 1 EAMCFMDPNLGPQVCYVQFPQ
Qy 635 CVFNRTAIYGYEPPKAKKPG
Db 61 CVFNRTAIYGYEPPKAKKPG
Qy 695 DIEEGVEGAGFDDKSVLMSOM
Db 121 DIEEGVEGAGFDDKSVLMSOM
Qy 755 VISCYEDKSWGTEIGWII
Db 181 VISCYEDKSWGTEIGWII

EBKRFQSPVFIASTLMENGLPEGTNSTSLIKE 358
EDILTGFKMHARGWRSVYCMKPAFAKGSAPIN 811
EDILTGFKMHARGWRSVYCMKPAFAKGSAPIN 418
YGYGRLKFLERFAINTIYPLTSLPLLYVCII 871
YGYGRLKFLERFAINTIYPLTSLPLLYVCII 478
ELSFATGILEMRWGSVGVDEWRNEQFWIGGI 931
ELSIATGVLELRWGSVGVDEWRNEQFWIGGI 538
KANDDEGDAELYFVKWTLLIPPTTILINMV 991
KAADDT-EFGELYFVKWTLLIPPTTILINMV 597
FWIVHLYPFLKGLMGRNQRTPTIIVLWVLLA 1051
FWIVHLYPFLKGLMGRNQRTPTIIVLWVLLA 657

900,237
344
9/15871
33
are 2689; DB 10; Length 506;
Ed. No. 4.2e-220;
Mismatches 1; Indels 0; Gaps 0;

EDRNDRYANRNTVFFDINLRLDGIQGVVYVGTG 634
EDRNDRYANRNTVFFDINLRLDGIQGVVYVGTG 60
CGGKKKSKRSRSDKKSKNKHVDSSVPVFNLE 694
CXGKKKSKRSRSDKKSKNKHVDSSVPVFNLE 120
RFGQSAAFVASTLMEYGGVQPSSTPESLLKEAII 754
RFGQSAAFVASTLMEYGGVQPSSTPESLLKEAII 180
ELTGFKMHARGWRSVYCMKPAFAKGSAPINUSD 814
ELTGFKMHARGWRSVYCMKPAFAKGSAPINUSD 240

QY 815 RLNQVLRWALGSVEILFSRHCPLWYGGGRKFLERFAYINTTIYPLTSLPLLVYCILPA 874
DB 241 RLNQVLRWALGSVEILFSRHCPLWYGGGRKFLERFAYINTTIYPLTSLPLLVYCILPA 300
QY 875 ICLLTGKFIMPEISNLSIWFIALFLSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISA 934
DB 301 ICLLTGKFIMPEISNLSIWFIALFLSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISA 360
QY 935 HLFVAVFQGLLKVLAGIDTNTFTVTSKANDEEGDFAELYPFKWTTLLIPPTTILIIINMGVV 994
DB 361 HLFVAVFQGLLKVLAGIDTNTFTVTSKANDEEGDFAELYPFKWTTLLIPPTTILIIINMGVV 420
QY 995 AGTSVAINSYGQSWGPLEGKLFPAFWIVHLYPFLKGLMGRONRTPTIVIVWAVLLASIF 1054
DB 421 AGTSVAINSYGQSWGPLEGKLFPAFWIVHLYPFLKGLMGRONRTPTIVIVWAVLLASIF 480
QY 1055 SLLWVRVDPFTTRLAGPNIQTGINC 1080
DB 481 SLLWVRVDPFTTRLAGPNIQTGINC 506

Search completed: February 19, 2003, 16:41:44
Job time : 39 secs

